UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

JUMIO CORPORATION,
Petitioner,

v.

FACETEC, INC.,
Patent Owner.

Patent Nos. 10,776,471; 11,157,606; 11,693,938; 11,874,910

Inter Partes Review Nos. IPR2025-00106, -00107, -00108, -00109

REPLY DECLARATION OF DR. CHRISTOPHER DAFT

Table of Contents

Page

I. INTRODUCTION 1

II. LEVEL OF ORDINARY SKILL IN THE ART 2

III. '471 PATENT 5

A. Derakhshani And Tanii (Ground 1A) 5

1. Derakhshani and Tanii teach "authenticating three-dimensionality of a user" ('471 Patent, Elements 1[pre] and 10[pre]) 5

a. Derakhshani 5

b. Tanii 16

2. Derakhshani and Tanii teach "comparing the first data to the second data to determine whether expected differences exist" ('471 Patent, Elements 1[d6] and 10[f]) and "when differences between the first data and the second data have expected distortion" ('471 Patent, Elements 1[d7] and 10[g]) 18

3. A POSITA would have been motivated to combine Derakhshani and Tanii 21

4. Derakhshani and Tanii teach "interpolating the first data and the second data to obtain estimated intermediate data" ('471 Patent, Elements 2[a] and 11[a]) 25

5. Derakhshani and Tanii teach "verifying the presence of the user's ears in the at least one first image, and verifying the absence or reduced visibility of the user's ears in the at least one second image" ('471 Patent, Claims 3 and 12) 26

a. Derakhshani 26

b. Tanii 28

6. Derakhshani and Tanii teach "comparing the first data, second data, or both to enrollment data" ('471 Patent, Element 5[a]) .29

7. Derakhshani and Tanii teach "the first data and the second data comprise biometric data" ('471 Patent, Claims 7 and 16) and "the first data and the second data comprise a mapping of facial features" ('471 Patent, Claims 8 and 17) 31

8. Derakhshani and Tanii teach "the first image and the second

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

image is of the user's face" ('471 Patent, Claim 9).....33

9. Derakhshani and Tanii teach "the user's face is held steady" ('471 Patent, Claim 19).....34

B. Derakhshani, Tanii, And Tahk (Ground 1B).....34

C. Derakhshani, Tanii, And Hoyos (Ground 1C).....34

D. Zhang And Tanii (Ground 2A).....36

1. Zhang and Tanii teach "comparing the first data to the second data to determine whether expected differences exist between the first data and the second data" ('471 Patent, Elements 1[d6] and 10[f]) and "differences between the first data and the second data have expected distortion" ('471 Patent, Elements 1[d7] and 10[g]).....36

a. Zhang.....36

b. Tanii.....48

2. A POSITA would have been motivated to combine Zhang and Tanii.....48

3. Zhang and Tanii teach "interpolating the first data and the second data to obtain estimated intermediate data" ('471 Patent, Elements 2[a] and 11[a]).....50

4. Zhang and Tanii teach "verifying the presence of the user's ears in the at least one first image, and verifying the absence or reduced visibility of the user's ears in the at least one second image" ('471 Patent, Claims 3 And 12).....51

5. Zhang and Tanii teach "the first data and the second data comprise biometric data" ('471 Patent, Claims 7 and 16).....51

E. Zhang, Tanii, And Hoyos (Ground 2C).....52

F. Tanii Is Analogous Art.....52

1. Tanii is from the same field of endeavor.....52

2. Tanii is reasonably pertinent to the problems faced by the inventors of the '471 Patent.....53

IV. '606 PATENT.....55

A. Derakhshani And Tanii (Ground 1A).....55

1. Derakhshani and Tanii teach "verifying three-dimensionality of a user's face using images of the user's face" ('606 Patent,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Elements 1[pre], 10[pre], and 19[pre]) 55
a. Derakhshani 55
a. Tanii 62
2. Derakhshani and Tanii teach "biometric data" ('606 Patent, Elements 1[b], 10[c], and 19[a]) 63
3. Derakhshani and Tanii teach "comparing[/compare] the first biometric data to[/with] second biometric data to determine whether differences between the at least one first image and the at least one second image match expected differences" ('606 Patent, Elements 1[f], 10[g], and 19[d]) 65
4. A POSITA would have been motivated to combine Derakhshani and Tanii 68
5. Derakhshani and Tanii teach "interpolating the first biometric data and the second biometric data to obtain estimated intermediate biometric data" ('606 Patent, Elements 2[a] and 11[a]) 72
6. Derakhshani and Tanii teach "verifying the presence of the user's ears in the at least one first image, and verifying the absence or reduced visibility of the user's ears in the at least one second image" ('606 Patent, Claims 3 and 12).....73
a. Derakhshani 73
b. Tanii 75
7. Derakhshani and Tanii teach "determining that the user's face is three-dimensional occurs at a server" ('606 Patent, Claim 14) 76
B. Derakhshani, Tanii, And Tahk (Ground 1B) 77
C. Zhang And Tanii (Ground 2A).....77
1. Zhang and Tanii teach "biometric data" ('606 Patent, Elements 1[b], 10[c], and 19[a]) 77
2. Zhang and Tanii teach "comparing[/compare] the first biometric data to[/with] second biometric data to determine whether differences between the at least one first image and the at least one second image match expected differences" ('606 Patent, Elements 1[f], 10[g], and 19[d]) 78
a. Zhang 78
b. Tanii 85

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

3. A POSITA would have been motivated to combine Zhang and Tanii 85

4. Zhang and Tanii teach "interpolating the first biometric data and the second biometric data to obtain estimated intermediate data" ('606 Patent, Elements 2[a] and 11[a]) 87

5. Zhang and Tanii teach "verifying the presence of the user's ears in the at least one first image, and verifying the absence or reduced visibility of the user's ears in the at least one second image" ('606 Patent, Claims 3 and 12) 88

D. Tanii Is Analogous Art 88

1. Tanii is from the same field of endeavor 88

2. Tanii is reasonably pertinent to the problems faced by the inventors of the '606 Patent 89

V. '938 PATENT 91

A. Derakhshani And Tanii (Ground 1A) 91

1. Derakhshani and Tanii teach "a human face" ('938 Patent, Element 1[pre]), "user's face" ('938 Patent, Elements 8[pre]/20[a]), and "a head of the person" ('938 Patent, Element 17[a]) 91

a. Derakhshani 91

b. Tanii 98

2. Derakhshani and Tanii teach "biometric data" ('938 Patent, Elements 1[a] and 17[d]) 99

3. Derakhshani and Tanii teach "the second biometric data does not have expected differences compared to the first biometric data, … the expected differences result from fish-eye type distortion" ('938 Patent, Element 1[d]), "when the first data and the second data do not have expected differences indicating the user exhibits three-dimensionality" ('938 Patent, Element 8[g]), "determine whether expected differences are not present" ('938 Patent, Element 17[d]), and "determining the user does not exhibit the expected degree of three-dimensionality" ('938 Patent, Element 20[h]) 101

4. A POSITA would have been motivated to combine Derakhshani and Tanii 104

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

5. Derakhshani and Tanii teach "compar[ing] at least portions of the first data, second data, or both to enrollment data" ('938 Patent, Element 5[a]) 108

6. Derakhshani and Tanii teach "the first biometric data and the second biometric data comprise [at least in part] image data of facial features" ('938 Patent, Claims 7 and 24) 110

B. Zhang And Tanii (Ground 2A) 110

1. Zhang and Tanii teach "biometric data" ('938 Patent, Elements 1[a] and 17[d]) 110

2. Zhang and Tanii teach "the second biometric data does not have expected differences compared to the first biometric data, … the expected differences result from fish-eye type distortion" ('938 Patent, Element 1[d]), "when the first data and the second data do not have expected differences indicating the user exhibits three-dimensionality" ('938 Patent, Element 8[g]), "determine whether expected differences are not present" ('938 Patent, Element 17[d]), and "determining the user does not exhibit the expected degree of three-dimensionality" ('938 Patent, Element 20[h]) 111

a. Zhang 111

b. Tanii 118

3. A POSITA would have been motivated to combine Zhang and Tanii 118

C. Tanii Is Analogous Art 120

1. Tanii is from the same field of endeavor 120

2. Tanii is reasonably pertinent to the problems faced by the inventors of the '938 Patent 121

VI. '910 PATENT 123

A. Derakhshani And Tanii (Ground 1A) 123

1. Derakhshani and Tanii teach "verifying[/evaluating] three-dimensionality of a[/the] user" ('910 Patent, Elements 1[pre], 10[pre], and 22[pre]) 123

a. Derakhshani 123

b. Tanii 130

2. Derakhshani and Tanii teach "comparing the first data to the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

second data to determine whether expected differences exist" ('910 Patent, Elements 1[d5] and 10[f]) and "comparing one or more aspects of the user's head … to determine whether expected differences, between the first image and the second image, exist" ('910 Patent, Element 22[d]) 132

3. A POSITA would have been motivated to combine Derakhshani and Tanii 135

4. Derakhshani and Tanii teach "verifying the presence of one or more features[/ears] … in the at least one first image, and verifying the absence or reduced visibility of the one or more features [/ears] … in the at least one second image" ('910 Patent, Claims 3 and 12) 139

a. Derakhshani 139

b. Tanii 141

5. Derakhshani and Tanii teach "comparing the first data, second data, or both to enrollment data" ('910 Patent, Element 5[a])142

6. Derakhshani and Tanii teach "biometric data" ('910 Patent, Claims 7 and 16) and "the first data and the second data comprise a map[ping] of facial features" ('910 Patent, Claims 8 and 17).....144

7. Derakhshani and Tanii teach "reflection of the illumination from a face of the user" ('910 Patent, Claim 18), "face of the user is held steady" ('910 Patent, Claim 19), and "user's head is the user's face" ('910 Patent, Claim 24).....146

B. Zhang And Tanii (Ground 2A).....147

1. Zhang and Tanii teach "comparing the first data to the second data to determine whether expected differences exist" ('910 Patent, Elements 1[d5] and 10[f]) and "comparing one or more aspects of the user's head … to determine whether expected differences, between the first image and the second image, exist" ('910 Patent, Element 22[d]) 147

a. Zhang 147

b. Tanii 154

2. A POSITA would have been motivated to combine Zhang and Tanii 154

3. Zhang and Tanii teach "verifying the presence of one or more

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

features[/ears] … in the at least one first image, and verifying the absence or reduced visibility of the one or more features [/ears] … in the at least one second image" ('910 Patent, Claims 3 and 12) 156

4. Zhang and Tanii teach "biometric data" ('910 Patent, Claims 7 and 16) 156

C. Zhang, Tanii, And Hoyos (Ground 2C) 157

D. Tanii Is Analogous Art 159

1. Tanii is from the same field of endeavor 159

2. Tanii is reasonably pertinent to the problems faced by the inventors of the '910 Patent 159

VII. OBJECTIVE EVIDENCE OF NONOBVIOUSNESS 161

A. The Challenged Claims Are Not Coextensive With FaceTec's ZoOm Product 161

1. Control of the camera 169

2. UI 174

3. Data encryption 179

4. Liveness checks 182

5. Removal of liveness data 188

B. FaceTec's Evidence Lacks Nexus In Fact To The Claimed Invention 192

1. Alleged praise lacks nexus to challenged claims, and Jumio was not a competitor 192

2. No proof iProov or Jumio copied the claimed invention 200

a. iProov 200

b. Petitioner 203

3. Claimed invention did not, alone, satisfy the long felt need, and there is no evidence of failure by others 205

4. Industry was not skeptical of the claimed invention 208

5. No nexus between alleged commercial success and claimed invention, and FaceTec statements are contradictory 210

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Petitioner's Updated Exhibit List

Exhibit Description	Exhibit #
U.S. Patent No. 10,776,471 ("the '471 Patent")	1001
File History of the '471 Patent	1002
Declaration of Dr. Chris Daft	1003
Dr. Chris Daft Curriculum Vitae	1004
U.S. Patent 8,437,513 ("Derakhshani")	1005
U.S. Patent Application Publication 2011/0299741 ("Zhang")	1006
U.S. Patent Application Publication 2002/0113884 ("Tanii")	1007
U.S. Patent Application Publication 2014/0028823 ("Tahk")	1008
U.S. Patent Application Publication 2004/0239799 ("Suzuki")	1009
U.S. Patent Application Publication 2010/0014720 ("Hoyos")	1010
U.S. Patent 9,077,891	1011
U.S. Patent 8,965,064	1012
Zhengyou Zhang et al., *A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry*, Institut National De Recherche En Informatique Et En Automatique (May 1994)	1013
U.S. Patent Application Publication 2010/0158319	1014

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Exhibit Description	Exhibit #
U.S. Patent Application Publication 2007/0127787	1015
U.S. Patent Application Publication 2004/0036574	1016
Eugene Hect & Alfred Zajac, *Optics*, Addison-Wesley Publishing Co. (1974)	1017
Anil K. Jain et al., *Introduction to Biometrics*, Springer (2011)	1018
B. Honlinger & H.H. Nasse, *Distortion*, Zeiss (Oct. 2009)	1019
Hue and Hatchet, *What is the Focal Plane?*, https://hueandhatchet.com/what-is-the-focal-plane/	1020
Introduction to Astronomy, https://physics.weber.edu/palen/phsx1040/lectures/ldistcomp.html	1021
Daniel Baker, *Face distortion is not due to lens distortion*, (May 5, 2012), https://bakerdh.wordpress.com/2012/05/05/face-distortion-is-not-due-to-lens-distortion/	1022
Richard Hartley & Andrew Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge Univ. Press (2nd ed. 2011)	1023
Daniel Moreno & Gabriel Taubin, *Simple, Accurate, and Robust Projector-Camera Calibration*, Brown Univ., mesh.brown.edu/calibration/	1024
Serge Belongie, *CSE 252B: Computer Vision II, Lecture 4: Planar Scenes and Homography* (Apr. 7, 2004)	1025
Merriam-Webster's Collegiate Dictionary (11th ed. 2014)	1026
Laurenz Wiskott, *Face Recognition by Elastic Bunch Graph Matching* (1999)	1027

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Exhibit Description	Exhibit #
Axis Communications, *CCD and CMOS sensor technology: Technical White Paper* (2010)	1028
Jeff Meyer, *What is Depth of Field? How aperture, focal length and focus control sharpness*, Digital Camera World (July 17, 2013), https://prism.org.gg/wp-content/uploads/2016/07/Depth-of-Field.pdf	1029
Brian Klug, *Understanding Camera Optics & Smartphone Camera Trends, A Presentation by Brian Klug*, Anandtech (Feb. 22, 2013), https://www.anandtech.com/Show/Index/6777?cPage=2&all=False&sort=0&page=2&slug=understanding-camera-optics-smartphone-camera-trends	1030
Portrait mode now available on iPhone 7 Plus with iOS 10.1, Apple (Oct. 24, 2016)	1031
The Timeline of Evolution of the Camera from the 1600s to 21st Century, Capture.com (May 5, 2023), https://www.capture.com/blogs/insights/evolution-of-the-camera	1032
John Biggs, *A Nokia Camera Phone That's More Like a Camera With a Phone Attached*, The New York Times (Nov. 29, 2007), ttps://www.nytimes.com/2007/11/29/technology/personaltech/29phone.html	1033
U.S. Patent 7,412,081	1034
U.S. Patent 8,457,367	1035
U.S. Patent 8,260,008	1036
Zac Hall, *Over a decade of selfies, starting with iPhone 4*, 9to5Mac (July 1, 2021), https://9to5mac.com/2021/07/01/iphone-4-selfie-celebration/	1037

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Exhibit Description	Exhibit #
Duane C. Brown, *Decentering Distortion of Lenses*, Photogrammertic Engineering (1966)	1038
Complaint, *FaceTec, Inc. v. iProov Ltd.*, C.A. No. 2:21-cv-02252, ECF No. 1 (D.N.V. Dec. 28, 2021)	1039
First Amended Complaint, *FaceTec, Inc. v. iProov Ltd.*, C.A. No. 2:21-cv-02252, ECF No. 18 (D.N.V. Feb. 24, 2022)	1040
Minute Entry, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv- 03623, ECF No. 38 (N.D. Cal. Sept. 18, 2024)	1041
Opposition to Motion to Intervene, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 40 (N.D. Cal. Sept. 30, 2024)	1042
Request for Voluntary Dismissal Without Prejudice, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 45 (N.D. Cal. Oct. 10, 2024)	1043
Opposition to Request for Voluntary Dismissal Without Prejudice, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 49 (N.D. Cal. Oct. 18, 2024)	1044
Reply in Support of Request for Voluntary Dismissal Without Prejudice, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 51 (N.D. Cal. Oct. 25, 2024)	1045
Complaint, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 1 (N.D. Cal. June 14, 2024)	1046
Motion to Intervene, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 36 (N.D. Cal. Sept. 16, 2024)	1047
Order Denying Motion to Voluntarily Dismiss Without Prejudice, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623, ECF No. 52 (N.D. Cal. Nov. 1, 2024)	1048

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Exhibit Description	Exhibit #
Declaration of Jessica C. Kaiser	1049
Docket, *FaceTec, Inc. v. Jumio Corp.*, C.A. No. 24-cv-03623 (N.D. Cal.)	1050
Board Email of April 30, 2025, authorizing Motion to Withdraw	1051
Supporting Declaration of Kyle W.K. Mooney for *Pro Hac Vice* Admission ("Mooney Decl.")	1052
Proposed Protective Order (marked-up version)	1053
Proposed Protective Order (clean version)	1054
Email chain between Parties dated September 8, 2025	1055
Email chain between Parties dated September 17, 2025	1056
Copy of transcript from conference call with the Board dated October 8, 2025	1057
Handwritten document from Expert Dr. Matthew A. Turk	1058 **[NEW]**
Deposition transcript of Dr. Matthew A. Turk (PUBLIC)	1059 **[NEW]**
Deposition transcript of Dr. Matthew A. Turk (CONFIDENTIAL)	1059 **[NEW]**
Second Declaration of Dr. Chris Daft (PUBLIC)	1060 **[NEW]**

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Exhibit Description	Exhibit #
Second Declaration of Dr. Chris Daft (CONFIDENTIAL)	1060 **[NEW]**
Frequently Asked Questions, FaceTec (September 7, 2024), https://dev.facetec.com/faq	1061 **[NEW]**
Architecture: Features & Configuration Options, FaceTec (March 16, 2023), https://dev.facetec.com/configuration-options	1062 **[NEW]**
Case Study – Kzen Networks, Tel Aviv, 2020, FaceTec (April 26, 2020), https://www.facetec.com/FaceTec_Case_Study_ZenGo.pdf	1063 **[NEW]**
NIST FRVT-PAD Commentary, FaceTec (July 8, 2022), https://facetec.com/FaceTec_NIST_FRVT-PAD_Comments.pdf	1064 **[NEW]**
Liveness Detection Security Report, FaceTec (February 6, 2023), https://www.facetec.com/FaceTec_Liveness_Security_Report_Q2_2022.pdf	1065 **[NEW]**
FaceTec, *FaceTec, Inc. Awarded Two Additional Patents for Identity Verification Security,* PR Newswire (June 5, 2025), https://www.prnewswire.com/news-releases/facetec-inc-awarded-two-additional-patents-for-identity-verification-security-302473637.html	1066 **[NEW]**
Remote Identity Proofing Practices: Attack Scenarios and Countermeasures, PWC (January 26, 2022), https://facetec.com/ENISA_RFI_Remote_ID_Attacks_FaceTec_Countermeasures.pdf	1067 **[NEW]**

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Exhibit Description	Exhibit #
ISO 30107-3 PRESENTATION ATTACK DETECTION CONFIRMATION LETTERS, iBeta (September 27, 2020), https://www.ibeta.com/iso-30107-3-presentation-attack-detection-confirmation-letters/	1068 **[NEW]**
Application Penetration Testing, Praetorian (October 4, 2024), https://www.praetorian.com/wp-content/uploads/2024/06/Praetorian-DataSheet-Application-Penetration-Testing-new.pdf	1069 **[NEW]**
Confirmation Letters, BixeLab (January 2, 2026), https://bixelab.com/confirmation-letter/	1070 **[NEW]**
Chris Burst, *FaceTec triples scope and payout in Spoof Bounty to thwart biometric liveness attacks*, Biometric Update (February 1, 2023), https://www.biometricupdate.com/202302/facetec-triples-scope-and-payout-in-spoof-bounty-to-thwart-biometric-liveness-attacks	1071 **[NEW]**
3D vs. 2D Biometrics, Part 2: Jay Meier on Liveness Detection and Revolutionizing IDV, ID Tech (February 9, 2025), https://idtechwire.com/3d-vs-2d-biometrics-part-2-jay-meier-on-liveness-detection-and-revolutionizing-idv/	1072 **[NEW]**
Transcript of *3D vs. 2D Biometrics, Part 2: Jay Meier on Liveness Detection and Revolutionizing IDV* podcast	1073 **[NEW]**
U.S. Patent Application 17/939,867, published as U.S. Patent Pub. 2023/0073410	1074 **[NEW]**
U.S. Design Patent No. 1074689	1075 **[NEW]**

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Exhibit Description	Exhibit #
U.S. Design Patent No. 987653	1076 **[NEW]**
U.S. Patent No. 12,423,398	1077 **[NEW]**
U.S. Patent Application 17/966,639, published as U.S. Patent Pub. No. 2023/0031087	1078 **[NEW]**
U.S. Patent No. 12,130,900	1079 **[NEW]**
U.S. Patent Application 18/627,421, published as U.S. Patent Pub. No. 2024/0348603	1080 **[NEW]**
U.S. Patent No. 11,657,132	1081 **[NEW]**
Deposition transcript of Josh Rose (PUBLIC)	1082 **[NEW]**
Deposition transcript of Josh Rose (CONFIDENTIAL)	1082 **[NEW]**
Technical Glossary, FaceTec (January 20, 2025), https://dev.facetec.com/glossary.	1083 **[NEW]**

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Exhibit Description	Exhibit #
Audio of *3D vs. 2D Biometrics, Part 2: Jay Meier on Liveness Detection and Revolutionizing IDV podcast*	1084 **[NEW]**
Declaration of Mina Ching	1085 **[NEW]**
Internet Archive Wayback Exhibit A	1086 **[NEW]**
Internet Archive Wayback Exhibit B	1087 **[NEW]**

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

I, Christopher Daft, make this declaration in connection with the proceeding identified above.

I. INTRODUCTION

1. I have been retained by counsel for Jumio Corporation (collectively, "Jumio" or "Petitioner") as a technical expert in connection with the proceedings identified above. I submit this reply declaration in support of Jumio's Replies to FaceTec's Patent Owner Response (POR) in the *Inter Partes* Review of U.S. Patent Nos. 10,776,471 ("the '471 patent") (IPR2025-00106, EX1001), 11,157,606 ("the '606 patent") (IPR2025-00107, EX1001), 11,693,938 ("the '938 patent") (IPR2025-00108, EX1001); and 11,874,910 ("the '910 patent") (IPR2025-00109, EX1001.) I described my qualifications in detail in my opening declaration dated November 6, 2024, which I understand is Exhibit 1003 to these proceedings (November 6, 2024, Daft Decl. (EX1003).) I have reviewed the exhibits in the exhibits list above as well as the Challenged Patents and their file histories.

2. I am being compensated at my standard consulting rate. I am also being reimbursed for expenses that I incur during my work. My compensation does not depend on the results of my study and analysis, the substance of my opinions, or the outcome of any proceeding involving the Challenged Claims. I have no personal or financial interest in the outcome of these matters, or in any

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

litigation involving the ’471, ’606, ’938, or ’910 patents (“the Challenged Patents”).

3. I am being compensated at my standard consulting rate. I am also being reimbursed for expenses that I incur during my work. My compensation does not depend on the results of my study and analysis, the substance of my opinions, or the outcome of any proceeding involving the Challenged Claims. I have no personal or financial interest in the outcome of these matters, or in any litigation involving the Challenged Patents.

4. Although I am not a lawyer and will not provide any legal opinions, I have been advised to apply certain legal standards in forming my opinions. I set forth my understanding of those legal standards in my opening declaration. (EX1003, ¶¶21-36.)

5. I understand that FaceTec and Dr. Turk have argued that the prior art does not teach certain limitations of the Challenged Claims of the Challenged Patents. I offer the following opinions in response.

II. LEVEL OF ORDINARY SKILL IN THE ART

6. Dr. Turk disagrees with the Petition’s level of a person of ordinary skill in the art (POSITA). (EX2061, ¶38.) His reason is because the Challenged Patents require “facial biometrics for identity verification” and “three-dimensional computer vision.” (EX2061, ¶¶39, 42, 44.) He also argues that “optics,”

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

"computer vision," and "biometrics" are too broad. (EX2061, ¶¶40-41, 43.) I disagree with Dr. Turk's arguments. The requirement of specialized knowledge in "facial biometric security" and "three-dimensional computer vision" is narrower than what a POSITA would need to in order to understand the Challenged Patents. This is particularly true because the claims of the Challenged Patents do not recite "biometric security," "facial authentication," or "three-dimensional computer vision." Further, the specifications of the Challenged Patents do not mention "three-dimensional computer vision" or suggest the need to be familiar with "stereo matching, epipolar geometry, structure from motion, multi-view stereo, shape modeling and representation, [or] depth estimation," as Dr. Turk mentions. (EX2061, ¶46.) Indeed, the Challenged Patents refer to the field of invention as encompassing "biometric security." (*See*, *e.g.*, EX1001, 1:28 for the '471 patent). In my opinion, as I explained in my opening declaration, a person with a Bachelor's degree in electrical engineering, computer engineering, computer science, physics, or a related field, and two years' work experience related to biometrics, facial authentication, computer vision, and/or optics with significant academic and/or work experience in both software development and optics would be able to understand the Challenged Patents and relevant prior art. (EX1003, ¶¶129-132.)

7. Dr. Turk also disputes whether my experience is relevant to the Challenged Patents. (EX2061, ¶¶49-53.) But at the time of the Challenged Patents, I had over 25 years of experience with imaging systems exhibiting the same optical effects seen in the Challenged Patents. (EX2082, 71:2-16, 72:18-73.) Because of at least my experience at Siemens Healthcare (EX1004, 7) I had much more than two years of experience in computer vision at the time of the Challenged Patents. (EX2082, 73:12-20.) Modern medical imaging is tied to computer vision due to the need to automatically interpret medical images. I also had over two years experience in biometrics and biometric security at the time of the Challenged Patents. (EX2082, 73:21-75:8.) My biometrics and biometric security experience involved biometric fingerprint recognition for user authentication. I note that the Challenged Patents refer to fingerprint authentication in the "Related Art" section. (*See*, *e.g.*, EX1001, 1:33-56 for the '471 patent).

8. I disagree with Dr. Turk's argument that my experience with fingerprinting and ultrasonic imaging is not relevant because the physics involved in ultrasonic imaging, fingerprinting, and not similar to the physics for distance-induced distortion. (EX2061, ¶49.) Light and sound both spread out as waves, and the ripples they form are described by the same wave equation. The list of fundamental physics concepts which apply in both fields is long: examples include wavelength and frequency, speed, reflection (echoes or mirrors), refraction

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

(bending), and interference (waves adding and canceling). Ultrasonic medical imagers can be accurately compared to cameras and optics because the wave equation governs these imaging systems. For example, considerations of depth of field, F-stop and diffraction effects transfer precisely between the optical and acoustic fields.

III. '471 PATENT

A. Derakhshani And Tanii (Ground 1A)

1. Derakhshani and Tanii teach "authenticating three-dimensionality of a user" ('471 Patent, Elements 1[pre] and 10[pre])

a. Derakhshani

9. Derakhshani teaches biometric authentication that involves determining whether images depict a three-dimensional subject (user). (EX1005, Title, Abstract, 17:22-26 (describing different spatial metric values indicating a three-dimensional subject or a two-dimensional spoof attack).)

10. Derakhshani's authentication process 600 is depicted in the flowchart of Figure 6 below. The authentication process 600 involves obtaining one or more images (step 602) and then two parts: (1) determining a liveness score for accepting or rejecting the images (steps 604, 606, 608, 610, and 616), and (2) determining a match score to accept or reject the user (steps 620, 622, 624, 626, 628, 630, and 632).

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

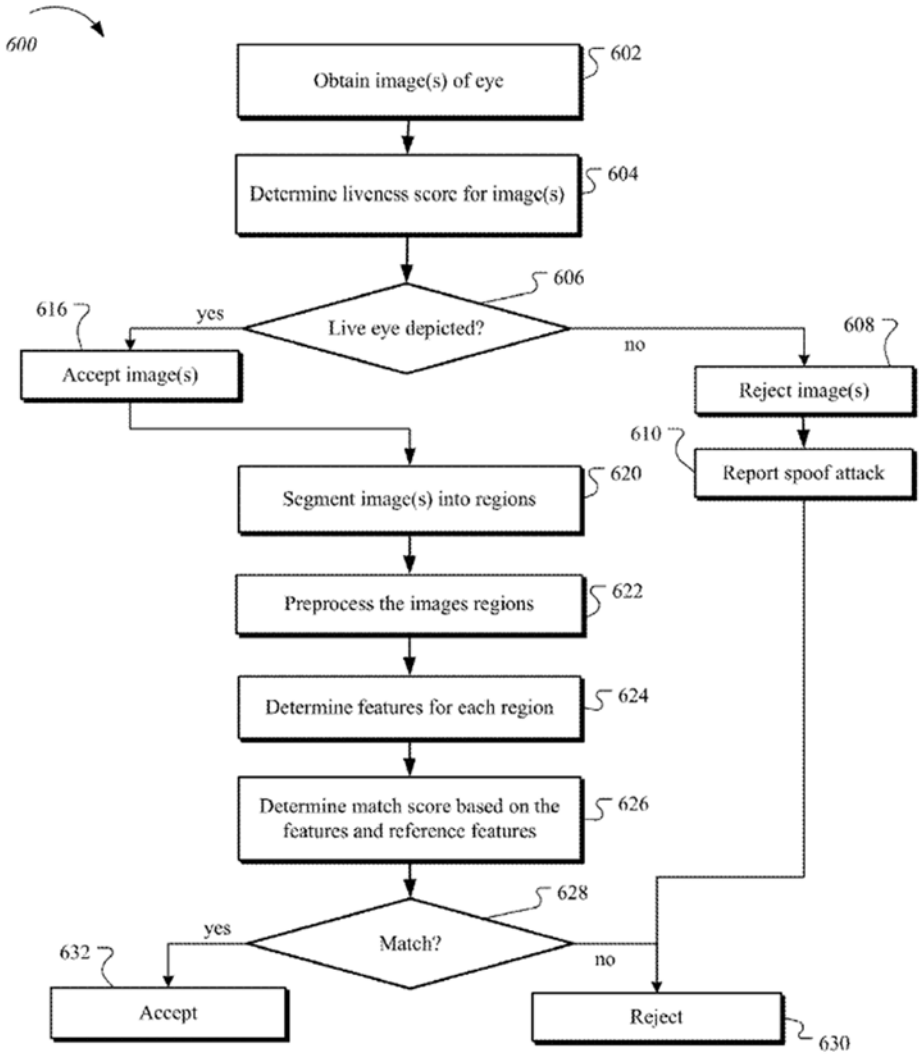


Fig. 6

11. <u>Liveness score</u>: The liveness score determined in step 604 represents the likelihood of a live eye in the captured image(s). (EX1005, 10:51-52, 11:5-8, 11:17-19.)

12. Figure 7 below shows the process 700 for determining the liveness score and provides further details for step 604 of Figure 6. (EX1005, 14:59-60.)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

The liveness score is determined based on a behavioral metric that measures eye movement (step 712), a spatial metric that measures three-dimensionality (step 714), and a reflectance metric that measures changes in reflectance (step 716). (EX1005, 15:27-29, 16:44-47, 17:45-47, 18:8-11.)

13. Spatial metric: The spatial metric measures three-dimensionality based differences in landmarks that appear in a plurality of captured images due to differences in the distance from the sensor to a landmark. (EX1005, 16:44-47, 17:22-26.) One image of the plurality of captured images may be a close-up image, and another image of the plurality of captured images may be captured at a different distance further away from the user. The close-up image may capture the eye vasculature (with enough precision), and the further-away image may capture the landmark(s) (with enough precision). Derakhshani names examples landmarks including "a nose, an ear," or a "mouth." (EX1005, 16:51-54, 17:16-17.) It also discusses "landmarks on a particular user's face," and a "landmark may be a portion of a face depicted in the images." (EX1005, 3:13-15, 17:29-36.)

14. Derakhshani discloses two approaches for determining the spatial metric: (1) a focus-distance approach, and (2) a parallax approach. Both approaches require a plurality of images captured with different distances from the sensor. The focus-distance approach "compar[es] the degree of focus for a landmark in images with different focus distances." (EX1005, 16:66-17:2.) The

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

parallax approach compares images where "landmarks within the images appear[] to move by different amounts because of differences in their distance from the sensor." (EX1005, 17:45-52.)

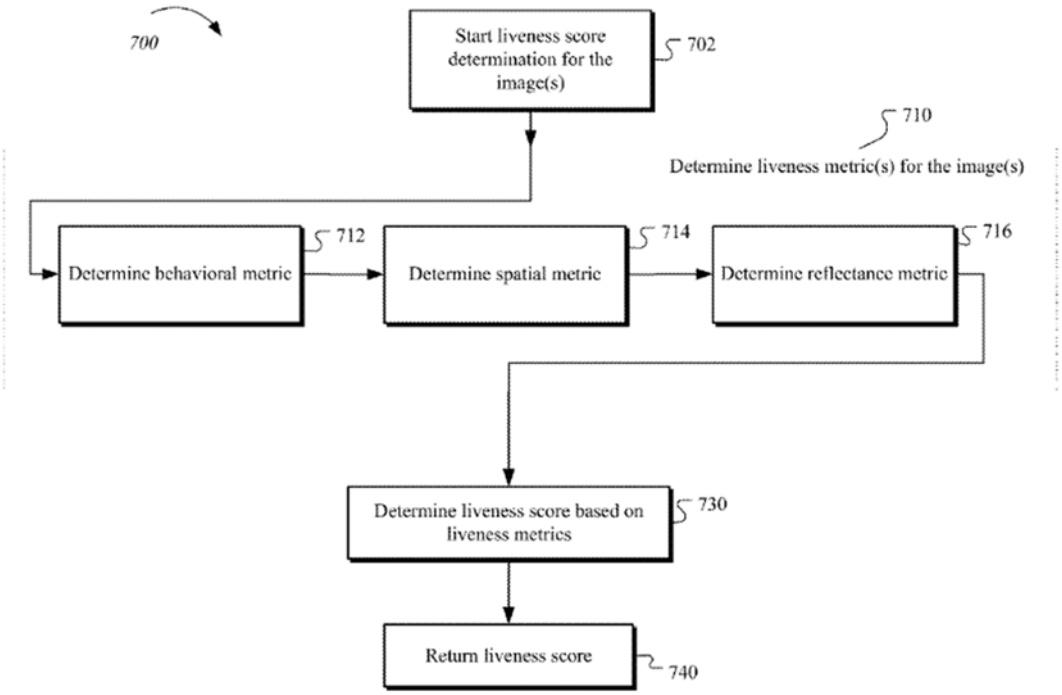


Fig. 7

15. Returning back to Figure 6, once the liveness score is determined, in step 606, the liveness score is compared to a threshold, and if the liveness score is low, then in steps 608 and 610, the images are rejected and a spoof attack is reported. (EX1005, 11:13-26.) If the liveness score indicates a live eye, then the images are accepted (step 616) and "subjected to further analysis to complete the authentication process." (EX1005, 11:42-45.) The "further analysis" refers to the match score part of the authentication process.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

16. Match score: In steps 620 and 622, regions of interest with the best views of the eye vasculature are identified in the images and preprocessed. (EX1005, 11:46-48, 12:23-24.) In step 624, the authentication process determines eye vasculature features for each image region. (EX1005, 13:9-11.) Then, in step 626, a match score is determined by comparing the features determined in step 624 and corresponding features from a reference record. (EX1005, 13:62-63.) The match score is compared to a threshold in step 628. The authentication process accepts (step 632) or rejects (step 630) the user based on whether or not the match score indicates a match of the captured image(s) to the reference record. (EX1005, 14:25-58.) Thus, Derakhshani authenticates the user based on a liveness of the user, including determining three-dimensionality via the spatial metric, and a match to a reference record.

17. Dr. Turk agreed that Derakhshani teaches authenticating three-dimensionality of a user (as elements 1[pre] and 10[pre] require). (EX1059, 128:21-129:1.) But he argues that Derakhshani does not teach capturing an image of the user's head or face. (EX2061, ¶¶77, 83.) However, none of the independent claims of the '471 patent requires capturing an image of the user's head or face, or authenticating the user using an image of the user's head or face. Rather, the independent claims of the '471 patent require authenticating three-dimensionality of a user, which, as I explained in my opening declaration, the Derakhshani-Tanii

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

combination teaches. (EX1003, ¶¶144-145 (citing EX1005, 1:11-25, 2:4-30, 3:14-15, 5:22-27, 6:3-5, 9:10-22, 16:44-18:4).)

18. Notably, Dr. Turk does not dispute that Derakhshani teaches capturing at least part of a user's face. He however argues that capturing part of the user's face (other than the eye) is incidental. (EX2061, ¶92.) I do not agree that capturing part of the user's face is incidental.

19. Derakhshani discloses determining a spatial metric "based on a distance from a sensor to a landmark that appears in a plurality of images each having a different respective focus distance." (EX1005, 16:44-47.) The determined spatial metric is used to determine three-dimensionality: "[a] high value for this spatial metric may indicate a three-dimensional subject and a higher likelihood that the subject is a live eye, while a low value may indicate a higher likelihood that the subject is a two-dimensional spoof attack." (EX1005, 17:22-26.)

20. Derakhshani specifically names "a nose, an ear," and a "mouth" as example landmarks identified in the captured images. (EX1005, 16:51-54, 17:16-17.) The passage at 17:29-36 discusses "landmarks on a particular user's face." (EX1005, 17:29-36.) And the passage at 3:13-15 states that the "landmark may be a portion of a face depicted in the images." (EX1005, 3:13-15.) Derakhshani also states that an image may comprise multiple "landmarks (e.g., an eye, nose, mouth,

and ear." (EX1005, 17:16-19; *see also* 3:5-7 ("two or more landmarks that appear in a plurality of the images"), 17:2-6 ("multiple landmarks").) Dr. Turk also agreed that Derakhshani captures an image having multiple landmarks. (EX1059, 136:25-137:3.) It would have been obvious for an image comprising an eye, nose, mouth, and ear, as Derakhshani discloses, to be an image of a face because these are facial features that when they are all included in an image, it would be an image of a user's face. (EX1005, 17:16-20.)

21. Derakhshani also names a "background object" as an example landmark. (EX1005, 16:51-54.) It would have been obvious for an image comprising an eye and a background object (landmark) to be an image of a user's face. For example, the background object may be a painting on a wall located behind the user. In order to capture the user's eye and the painting in the image, the sensor would have to be back far enough from the user such that the user's face would be captured. Thus, Dr. Turk's suggestion (at, *e.g.*, ¶¶85 and 89 of his declaration (EX2061)) that Derakhshani captures images of the user's eyes only is wrong. I also note that Derakhshani does not disclose capturing the user's eye vasculature as its primary purpose, as Dr. Turk argues at ¶¶87 and 92 of his declaration.

22. Derakhshani further states that its authentication is "based ***in part*** on one or more images of the white of an eye 410." (EX1005, 5:65-67.) Being based

"in part" means that some of the captured images are of the white of an eye 410 (such as the images captured using the sensor 6-12 cm from the eye (EX1005, 13:35-40)), and some of the captured images are not just of the white of an eye, such as images of the user's face.

23. Dr. Turk's declaration argument that the spatial metric determination relies on characteristics of the user's eye only is also incorrect. (EX2061, ¶86 (citing EX1005, 4:19-21).) Derakhshani's spatial metric involves capturing a plurality of images at difference distances. (EX1005, 16:44-47, 17:45-47.) One image of the plurality of captured images may be a close-up image, and another image of the plurality of captured images may be captured at a different distance further away from the user. The close-up image may capture the eye vasculature (with enough precision), and the further-away image may capture the landmark(s) (with enough precision).

24. The passage at 4:19-21 that Dr. Turk cites describes image(s) used for determining the match score. The match score is determined using image(s) that capture the vasculature of the eye and are "subjected to further analysis" (step 616) ***after and separate from*** the determination of the spatial metric (step 714, which is part of step 710 performed as step 604). (EX1005, 14:59-64, 16:44-18:7.) As Dr. Turk testified, the eye vasculature is determined after determining the liveness score and determining the liveness score is separate from determining the match

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

score. (EX1059, 130:3-7, 132:9-11.) Further, Dr. Turk agreed during his deposition that the spatial metric determination is not solely based on characteristics of the user's eye. (EX1059, 126:4-13.) Dr. Turk also agreed that Derakhshani captures images having more than the user's eye. (EX1059, 136:16-24.)

25. Further, a POSITA would have understood a prior art reference based on not only what is depicted in the figures, but also the teachings discussed in the specification. Even Dr. Turk looks to the specification when reviewing a prior art reference. (EX1059, 16:3-17:24.) Accordingly, Dr. Turk's declaration arguments that Derakhshani captures images of the eyes only because that is what is depicted in the figures is without merit because Derakhshani's specification suggests otherwise. (EX2061, ¶¶77, 87-88; EX1005, 16:44-54, 17:16-17.) Dr. Turk also testified that if there is an embodiment with the image including a nose, ear, mouth, and eye, it would be an image that is not depicted in Derakhshani figures. (EX1059, 137:4-17.) Indeed, Derakhshani discloses such an embodiment, and thus, its figures do not show all of the disclosed embodiments. (EX1005, 17:16-19.)

26. Dr. Turk argues that Derakhshani's system is not adapted to capture the user's face or head. (EX2061, ¶¶87, 89.) I disagree. To make an accurate determination of the spatial metric, a POSITA would have understood that there

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

must be a discernable difference in the degree of focus (for the focus-distance approach) or movement of the landmarks (for the parallax approach), which would mean there must be a discernable difference in distance between the user and the sensor. This would mean at least one captured image would be taken such that the sensor is far enough away from the user, thereby including more than just the user's eye in the image. Derakhshani discusses this at 17:55-59: "[i]f all the landmarks in the image undergo the ***same apparent displacement*** due to relative motion of the sensor, ***i.e., the difference in the parallax effect for the landmarks is small***, then the subject viewed by the camera has [a] higher likelihood of being a two-dimensional spoof attack." (EX1005, 17:55-59.) A POSITA would have understood from this passage that a large difference in the parallax effect for the landmarks would be needed when the object is three-dimensional in order to avoid a false determination of a two-dimensional spoof attack.

27. Further, Derakhshani teaches achieving a finer resolution by using a reduced dynamic range for the sensor. "[O]ne or more images may be captured using a reduced dynamic range for the sensor (e.g., a camera) so that a finer resolution in intensity of detected light is achieved in a range within which it occurs in the captured images. In this manner, the intensity or color scale can be zoomed in to reveal more subtle changes in the level of the detected image signal." (EX1005, 19:24-32.) Thereby, by adjusting the parameters, the same sensor may

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

be used to capture images of the vasculature of an eye with a finer resolution (lower dynamic range) and images of the face. This means Derakhshani's system is adapted to capture the user's head/face.

28. Dr. Turk also argues that simultaneous capture of the eye vasculature and face in the same image with enough precision for both is required. (EX2061, ¶91.) I disagree based on Derakhshani's disclosures of capturing a plurality of images for the spatial metric determination. For the focus-distance approach, the plurality of images have different respective focus distances for a given landmark. Derakhshani provides an example where the landmarks in an image include "an eye, nose, mouth, and ear." (EX1005, 17:16-17.) It explains that a landmark may be "in focus" or may be blurred (also described as "the degree of focus is low"). (EX1005, 16:56-66.) A POSITA would have understood Derakhshani's disclosures of different focus distances in different images and a landmark being in focus and blurry means the landmark is in focus in one image and blurred in another image. The spatial metric is determined based on distances of the landmarks. (EX1005, 16:66-17:6, 17:12-22, 17:49-52.) A high value for the spatial metric indicates a three-dimensional subject and a low value indicates a two-dimensional spoof attack. (EX1005, 17:22-26.) A POSITA would have understood that large estimated distances from the landmarks would occur when some of the landmarks are blurred (out of focus) and result in a high value for the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

spatial metric. Thus, Derakhshani does not need to capture images that simultaneously capture both the vasculature of an eye and a face with enough precision in order to determine three-dimensionality or authentication, as Dr. Turk argues. (EX2061, ¶91.)

29. Instead, a POSITA would have understood that Derakhshani may be able to determine three-dimensionality when not simultaneously capturing both the vasculature of an eye and face with enough precision. For example, Derakhshani may capture a close-up image with the user's eye in focus but the ear is not in focus, and may also capture another image that is further away from the user where the ear is in focus. The spatial metric may be determined using the close-up image and the further-away image, both having a discernable difference in distance of the ear landmark, thereby resulting in a high spatial metric and determination of three-dimensionality. The match score may be determined using the close-up image. In such a situation, an image that simultaneously captures both the eye vasculature and the face with enough precision is not required. Thus, not every image needs to be close enough to capture the vasculature of the eye with sufficient clarity, as Dr. Turk argues. (EX2061, ¶¶78, 82.)

b. Tanii

30. Tanii also teaches capturing images of a user's three-dimensional face involving distance-induced distortions, as shown in Figures 3A-3B and 4A-4B

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

(reproduced below). (EX1007, Abstract ("an image is captured with the user's face at a close range"), [0006], [0013], [0047]-[0049], [0061]-[0063].) Dr. Turk does not dispute this, but rather argues that Tanii does not identify a face. (EX2061, ¶96.) This is incorrect. Tanii states that it "detects the size of the image of the main object." (EX1007, [0074].) Tanii explains that "the image of the main object, particularly the image of the face of the user who is holding the cellular phone 1, is captured." (EX1007, [0061].) Thus, Tanii detects the size of the user's face. In order to detect the size of the user's face, it must first detect the user's face. Further, Tanii describes different ways to detect the user's face, such as "extracting the contours of the main object using image processing." (EX1007, [0069].) Tanii also describes "identif[ying] the region of the image of the main object based on the location of clear edges in the image as well as on the color distribution in the image." (EX1007, [0076].)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Fig.3A

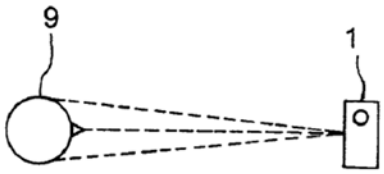


Fig.3B

Fig.4A

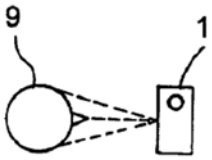
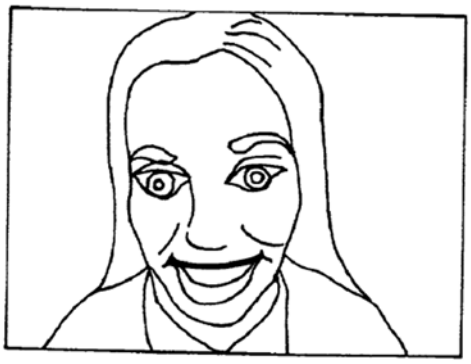


Fig.4B



2. Derakhshani and Tanii teach "comparing the first data to the second data to determine whether expected differences exist" ('471 Patent, Elements 1[d6] and 10[f]) and "when differences between the first data and the second data have expected distortion" ('471 Patent, Elements 1[d7] and 10[g])

31. Tanii teaches the "expected differences" and "expected distortion" of elements 1[d6] and 10[f], and elements 1[d7] and 10[g]. (EX1007, Figs. 3A-3B, 4A-4B, 5, [0047]-[0049].) Tanii explains that when there is a sufficient distance

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

between the main object 9 and the cellular phone 1, the image appears natural. (EX1007, [0047], Figs. 3A-3B.) When the main object 9 and cellular phone 1 are too close to each other, the perspective is exaggerated resulting in an unnatural image. (EX1007, [0047], Figs. 4A-4B.) This exaggeration in perspective and unnaturalness is an expected distortion because it occurs due to known and expected differences that result from the light ray striking different points of the main object 9. (EX1007, [0048]-[0049], Fig. 5.)

32. Dr. Turk does not dispute these teachings. Instead, he focuses his arguments on whether Tanii evaluates degree of distance-induced distortion, or measures, assesses, or undoes three-dimensional distortion. (EX2061, ¶¶80, 98.) But the '471 claims do not require evaluating the degree of distance-induced distortion, or measuring or undoing three-dimensional distortion. Further, the Derakhshani-Tanii combination relies on Derakhshani for its teachings of authenticating the user, which Dr. Turk does not dispute Derakhshani teaches.

33. Moreover, Dr. Turk is incorrect that Tanii does not assess three-dimensional distortion and that it assumes an image requires correction. (EX2061, ¶165.) Tanii determines the distance from the main object to the camera. (EX1007, [0101].) It discloses "where the distance to the main object equals or exceeds the threshold value D1, it is determined that the main object and the imaging portion 2 are located a sufficient distance apart and that as a result no

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

correction is needed." (EX1007, [0101].) That is, when there is no distance-induced distortion, Tanii does not perform a correction. Tanii also teaches determining the degree of distance-induced distortion and applying the appropriate amount of correction. It discloses "[w]here the distance to the main object is less than the threshold value D1" or "less than the threshold value D2," then "a low degree of correction" or "a high degree of correction" is needed. (EX1007, [0101].) It further discloses "[i]t can be automatically determined whether or not correction is needed." (EX1007, [0136], Claim 7.) In other words, Tanii identifies and assesses distance-induced distortions.

34. Dr. Turk also argues that Derakhshani compares the captured images to an expected plane/shape, but does not compare data of two captured images together. (EX2061, ¶84.) This is incorrect. Dr. Turk points to 3:1-6 and 17:27-36 to support his argument, but these passages describe one embodiment. There are other passages in Derakhshani, such as 16:51-54 and 17:45-47, that teach comparing data of two captured images together. The 16:51-54 passage and surrounding text describe the spatial metric being determined based on "a distance from a sensor to a landmark that ***appears in a plurality of the images*** each having a different respective focus distance." (EX1005, 16:44-54 (emphasis added).) It states that "the focus distance may be adjusted for different images." (*Id.*; *see also* 16:44-54 ("the plurality of the images with different focus distances").) The

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

17:45-47 passage and surrounding text describe "parallax of two or more landmarks that ***appear in a plurality of the images***," where the "plurality of images [are] taken from different perspectives on the subject" and "the sensor is moved about the subject during image capture." (EX1005, 17:45-64 (emphasis added).) Thus, Derakhshani captures a plurality of images having different distances, and the data from the images are compared to each other.

35. Dr. Turk further argues that Tanii aesthetically corrects images and states that its alleged "process of enlargement" does not make up for the deficiencies of Derakhshani and is incompatible with Derakhshani. (EX2061, ¶¶96, 98.) However, Dr. Turk misunderstands the Derakhshani-Tanii combination which does not use Tanii's aesthetic/warp correction. The Derakhshani-Tanii combination includes Tanii's teachings of distance-induced distortions in images of a user's face. (EX1003, ¶¶134-143.) Further, whether Tanii operates the same way when the image does not include a face is irrelevant because Tanii teaches capturing and detecting the user's face. (EX2061, ¶97; EX1007, [0062].)

3. A POSITA would have been motivated to combine Derakhshani and Tanii

36. Dr. Turk argues that there is no motivation to combine Derakhshani and Tanii because Tanii destroys Derakhshani's "basic objective" that requires distortions to operate. (EX2061, ¶¶162-163, 166.) Dr. Turk's arguments are however premised on his misunderstanding of the Derakhshani-Tanii combination

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

and the Petition's motivation to combine. The teachings of Tanii that are part of the Derakhshani-Tanii combination are its teachings of distance-induced distortions in images of a user's face—not teachings of image correction. (EX1003, ¶¶134-143; EX2061, ¶165.)

37. Additionally, Dr. Turk's argument that images at different distances having different distortion is not a motivation to combine again shows that he misunderstands the Derakhshani-Tanii combination. (EX2061, ¶¶99, 167.) As I explained in my opening declaration, Tanii teaches an obvious alternative to Derakhshani's focus-distance and parallax approaches. (EX1003, ¶¶134, 138.) Derakhshani's approaches require a camera with extreme sensitivity, not suitable for mobile phones that have wide angle lenses to capture a wide field of view. (EX1003, ¶139.) I note that Dr. Turk acknowledges that Derakhshani's approach has limited applicability. (EX2061, ¶168.) Dr. Turk also acknowledges that Tanii's teachings apply in situations where images are captured using a wide-angle lens. (EX2061, ¶169.) Thus, a POSITA would have been motivated to use Tanii's method when using a mobile device that has a wide-angle lens to capture the user's image(s). (EX1003, ¶¶140-143.) Dr. Turk fails to rebut this motivation.

38. I disagree with Dr. Turk's argument that Derakhshani and Tanii are incompatible. (EX2061, ¶169.) Dr. Turk's argues that this alleged incompatibility is due to Derakhshani and Tanii having different imaging requirements. (EX2061,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

¶169.) Dr. Turk argues that Derakhshani needs to capture the eye vasculature with sufficient clarity and the camera must be positioned very close to the user's eye, eliminating the ability to capture the user's face or the ability to use Tanii's correction method designed for capturing images at an arms length. This argument misunderstands Derakhshani's disclosures and the Derakhshani-Tanii combination. Derakhshani's system captures a plurality of images at different conditions. As I explained above, in Derakhshani, different images may be used by different functions of its authentication process. For the spatial metric determination (part of the liveness detection), Derakhshani uses images captured at different focus distances, or from different perspectives to analyze landmarks such as "a nose, an ear," "a mouth," or a "background object." (EX1005, 16:51-54, 17:16-19.) A POSITA would have understood that capturing an image containing facial landmarks results in capturing an image of the user's face. For the match score determination (eye vasculature comparison), Derakhshani uses close-up images to resolve the eye vasculature with sufficient precision. (EX1005, 13:35-40.) A POSITA would have understood that Derakhshani's multi-image framework accommodates both types of captured mages: some images optimized for the spatial metric determination (which may include facial features at varying distances from the sensor), and some images optimized for vasculature matching (which may include images of the eye vasculature). Dr. Turk assumes that all

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

images must simultaneously capture both the eye vasculature and the facial landmarks with precision, but Derakhshani does not disclose or suggest such a requirement. Further, Dr. Turk's argument about incompatibility supports the motivation to combine because Derakhshani's approach is limited and applying Tanii's teachings would result in improving the implementation of biometric authentication in a mobile device as Derakhshani already envisions. (EX1005, 5:23-26; EX1003, ¶¶138-143.) Derakhshani's focus-distance and parallax approaches use specialized camera capabilities that may not be suitable for mobile phones with wide-angle lenses. Tanii teaches capturing images of the user's face using a cellular phone with a wide-angle lens at an arms length. (EX1007, [0007], [0047].) Tanii discloses that when "the user's face is close to the camera" with "a wide-angle focal length," "the exaggeration warp becomes more noticeable." (EX1007, [0007].) As I explained my opening declaration, a POSITA would have wanted to implement Derakhshani's authentication process on a mobile device, and would have been motivated to use Tanii's teachings about distance-induced distortion of an image of a user's face as an alternative or supplemental way to determining spatial metric. Derakhshani teaches using "a reduced dynamic range for the sensor . . . so that a finer resolution in intensity of detected light is achieved," demonstrating that the same sensor can be adjusted to capture different types of images by modifying camera parameters. (EX1005, 19:24-32.) A

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

POSITA would have understood that similar adjustments would allow Derakhshani's system to capture both close-up images for the match score determination and further away images for the spatial metric determination (using Derakhshani's and/or Tanii's approach). As I explained in my opening declaration, a POSITA would have known how to apply Tanii's teachings to Derakhshani's system using well-known optics principles relating to camera design, and there would have been a reasonable expectation of success. (EX1003, ¶141.) Dr. Turk does not dispute this.

39. Further, I disagree with Dr. Turk's argument that capturing images of the user's face would render Derakhshani's system inoperable for its intended purpose. (EX2061, ¶¶95, 162.) As I explained above, Derakhshani's system captures images of the user's face for determining the spatial metric, thereby showing it is operable when capturing images of the user's face.

4. Derakhshani and Tanii teach "interpolating the first data and the second data to obtain estimated intermediate data" ('471 Patent, Elements 2[a] and 11[a])

40. Dr. Turk argues that Derakhshani does not teach elements 2[a] and 11[a] because it does not generate a two-dimensional plane. (EX2061, ¶¶100-102.) However, Dr. Turk misunderstands Derakhshani's disclosures. Derakhshani states that "the locations of landmarks (e.g., an eye, nose, mouth, and ear) in the two dimensional image will occur in **a** two dimensional plane." (EX1005, 17:16-18.)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

"For example, the locations of multiple landmarks may be fit to ***the closest two dimensional plane***." (EX1005, 17:19-22 (emphasis added).) A POSITA would have understood that "a two dimensional plane" refers to any two-dimensional surface such as a two dimensional plane that intersects with the x-axis and y-axis at (5,5), a two-dimensional plane that intersects with the x-axis and the y-axis at (5,4), or a two-dimensional plane that intersects with the x-axis and the y-axis at (4,4). Derakhshani's fitting to the closest two dimensional plane means it determines the points of the two-dimensional plane for generating it. Thus, Dr. Turk is wrong that Derakhshani's interpolation does not generate any planes and that its landmarks are fit to an existing plane.

41. Dr. Turk also argues that there is no motivation to combine Derakhshani and Tanii. As discussed above, Dr. Turk's arguments are incorrect.

5. Derakhshani and Tanii teach "verifying the presence of the user's ears in the at least one first image, and verifying the absence or reduced visibility of the user's ears in the at least one second image" ('471 Patent, Claims 3 and 12)

a. Derakhshani

42. Dr. Turk acknowledges that Derakhshani's captures an image of an ear, but argues that it does not teach verifying the presence, absence, or reduced visibility of the ears. (EX2061, ¶¶103-104.) But Derakhshani states that "a landmark (e.g., an iris, an eye corner, a nose, ***an ear***, or a background object) ***may***

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

be identified and located in the plurality of images with different focus distances." (EX1005, 16:51-54 (emphasis added).) It also states that "determining 714 the spatial metric comprises parallax of two or more landmarks that ***appear in a plurality of the images***. . . . This ***parallax effect may be measured*** and used as a spatial metric that reflects the three-dimensional nature of a subject in the view of the sensor." (EX1005, 17:45-55 (emphasis added).) Identifying and locating a landmark or measuring parallax for a landmark that appears in the images means verifying the presence of a landmark, including verifying the presence of an ear in an image. And just like how capturing an image of the user's face is not incidental (see discussion above), capturing an image of the ear is not incidental, as Dr. Turk argues. (EX2061, ¶¶103-104.)

43. Dr. Turk also argues that Derakhshani's use of landmarks for determining the spatial metric undercuts the idea of an ear being absent or having reduced visibility. (EX2061, ¶104.) I disagree. Derakhshani discloses different degrees of focus (a low degree of focus or a high degree of focus) for a landmark. (EX1005, 16:60-17:11.) A landmark having a low degree of focus would be blurry, thereby having reduced visibility. Dr. Turk testified similarly, but used the terms "resolution" and "definition." (EX1059, 142:12-143:21.) Derakhshani explains that "[w]hen the degree is low for a landmark, smaller higher frequency components are expected. By comparing the degree of focus for a landmark in

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

images with different focus distances, the distance from the sensor to the landmark may be estimated." (EX1005, 16:65-17:2.) The distances, including the distance for a blurry landmark, are then used to determine the spatial metric. (EX1005, 17:19-26.) The use of distance information for a landmark shows that Dr. Turk is wrong that Derakhshani merely uses a landmark to estimate the degree of focus. (EX2061, ¶105.) The focus-distance approach involves capturing a first image with a landmark in focus and a second image with the same landmark not in focus (blurry). The spatial metric determination involves comparing the degrees of focus in the first and second images. Dr. Turk confirmed this understanding. (EX1059, 144:11-16, 145:9-12.) Thus, Derakhshani teaches determining the spatial metric of an image of a user's ear having reduced visibility.

b. Tanii

44. Dr. Turk argues that nothing in Tanii suggests verifying the presence or absence of ears in an image. (EX2061, ¶106.) But Tanii's images show otherwise. Figure 3B (below) shows the presence of the user's ear in the first image, and Figure 4B (below) shows the absence of the user's ear in the second image. Further, Dr. Turk's focus on whether Tanii verifies the presence, reduced visibility, or absence of the ears shows his misunderstanding of the Derakhshani-Tanii combination. (EX2061, ¶106.) As discussed above, Derakhshani teaches verifying the user's ears.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Fig.3A

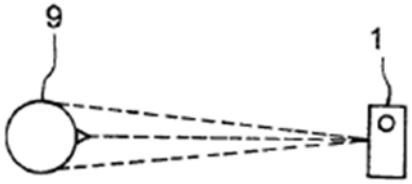


Fig.3B

Fig.4A

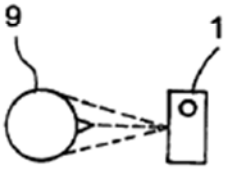
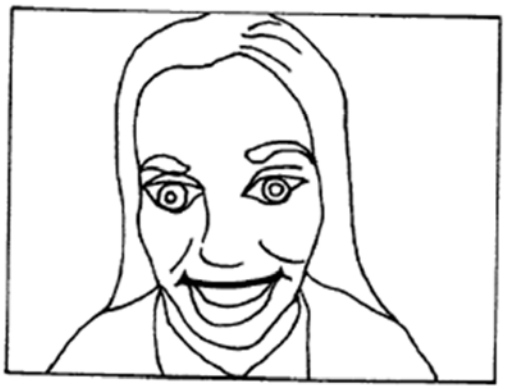


Fig.4B



6. Derakhshani and Tanii teach "comparing the first data, second data, or both to enrollment data" ('471 Patent, Element 5[a])

45. Derakhshani teaches identifying landmarks and using the "relative positions of landmarks on a particular user's face . . . [to] generate a three-dimensional model that is stored as part of a reference record." (EX1005, 17:32-

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

36.) The features in the image (e.g., of the eye) are extracted and compared to the reference record. (EX1005, 4:19-24, 7:19-34, 9:59-67, 13:62-14:9.) In other words, Derakhshani uses both terms "landmarks" and "features" to describe the data of the reference record, thereby Derakhshani does not make a distinction between "landmarks" and "features," as Dr. Turk argues. (EX2061, ¶¶112-116.) And "landmarks" are also used for the spatial metric determination, thereby the data for liveness detection is the same as the data used for the reference record (what Dr. Turk refers to as "enrollment records" (EX2061, ¶116)).

46. Dr. Turk's argument that "landmarks" and "features" not synonymous because the "landmarks" used for liveness determination are separate from the "features" used for authentication is incorrect. (EX2061, ¶¶115-116.) As discussed above, Derakhshani determines the liveness score as part of the authentication process, and the spatial metric is determined as part of determining the liveness score. The passage at 8:60-67 explains that the liveness score is determined as part of Derakhshani's authentication process: "authentication module 540 may receive image data from the requesting user's computing device … and may process the image data to authenticate the user. In some implementations, the ***authentication module may determine liveness scores*** for images received from a user and may accept or reject the images based on the liveness scores." (EX1005, 8:60-67 (emphasis added).) Also, the passage at

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

14:42-45 states "[i]f the ***liveness score*** indicates a high likelihood that live eye is depicted in the images, the one or more images are accepted 616 and subjected to further analysis to ***complete the authentication process.***" "complet[ing] the authentication process" means the liveness score is part of the authentication process. Further, Derakhshani names the eye as a landmark, and the eye is also a feature. (EX1005, 17:16-17 ("***landmarks (e.g., an eye,*** nose, mouth, and ear).)

7. Derakhshani and Tanii teach "the first data and the second data comprise biometric data" ('471 Patent, Claims 7 and 16) and "the first data and the second data comprise a mapping of facial features" ('471 Patent, Claims 8 and 17)

47. Dr. Turk argues that Derakhshani does not teach claims 7, 8, 16, and 17 of the '471 patent because its landmarks are not "biometric data." Dr. Turk's first reason is because Derakhshani teaches "image landmarks," which may be a background object. (EX2061, ¶118.) But this is incorrect. First, I disagree with Dr. Turk's separation of landmarks into what he calls "image landmarks" and "anatomical landmarks." (EX2061, ¶58.) Dr. Turk seems to create this separation because "image landmarks" are used for the liveness score determination and "anatomical landmarks" are used for the match score determination. (EX2061, ¶58.) However, Derakhshani teaches the eye as being an example landmark used for both the liveness score determination and the match score determination, thereby showing that a landmark cannot be classified as either an "image

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

landmark" or an "anatomical landmark." (EX1005, 14:25-29; 17:16-19.) Further, even if a background object does not contain "biometric data," this does not mean all "image landmarks" cannot contain biometric data. Dr. Turk even testified that his meaning of "biometric data" is "data about a human, human biology or body or something." (EX1059, 44:4-13.) Dr. Turk cites 15:51-54 of Derakhshani, which provides "an iris, an eye corner, a nose, an ear, or a background object" as example landmarks. (EX2061, ¶112 (citing EX1005, 15:51-54).) Because an image of an iris, an eye corner, a nose, and an ear may have data about human biology or body, then Derakhshani's so-called landmarks may contain "biometric data."

48. Dr. Turk's second reason is because Derakhshani's landmark may be an iris, eye corner, or ear, which is not used to identify a user. (EX2061, ¶118.) But Derakhshani uses the iris, eye corner, or ear to determine a liveness score or a match score, both of which are used to authenticate the user. (EX1005, 11:46-51, 16:44-54, 17:12-26, 19:16-18, 21:8-28, Figs. 6-8.)

49. I note that Dr. Turk's requirement provided in his declaration that "biometric data" be used to identify a user is inconsistent with the '471 specification, which states: "[t]he term biometric data may include among other information biometric information concerning facial features and ***path parameters***. ***Examples of path parameters may include an acceleration and speed of the mobile device, angle of the mobile device during image capture, distance of the***

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

mobile device to the user, path direction in relation to the user's face position in relation to the user, or any other type parameter associated with movement of the mobile device or the user face in relation to a camera. Other data may also be included such as GPS data, device identification information, and the like." ('471 patent (EX1001), 6:37-47.) This is inconsistent with his testimony where he agreed with the specification's description of "biometric data." (EX1059, 50:11-14.)

50. Dr. Turk's third reason is because Derakhshani's landmarks are not a mapping of facial features since it does not capture images of the user's face. (EX2061, ¶119.) As I discussed above, Derakhshani does capture images of faces, the images may include facial feature landmarks such as "an eye, nose, mouth, and ear." (EX1005, 17:16-19.)

8. Derakhshani and Tanii teach "the first image and the second image is of the user's face" ('471 Patent, Claim 9)

51. Dr. Turk's only argument why Derakhshani-Tanii does not teach claim 9 is because Derakhshani captures images of the eyes and not the face. (EX2061, ¶94.) As I discussed above, Dr. Turk is wrong.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

9. Derakhshani and Tanii teach "the user's face is held steady" ('471 Patent, Claim 19)

52. Dr. Turk's only argument for why Derakhshani-Tanii does not teach claim 19 is because Derakhshani captures images of the eyes and not the face. (EX2061, ¶94.) As I discussed above, Dr. Turk is wrong.

B. Derakhshani, Tanii, And Tahk (Ground 1B)

53. Ground 1B challenges claim 14, which requires "one or more prompts are ovals on the screen within which the face of the user is placed."

54. Dr. Turk's only argument why Derakhshani-Tanii-Tahk does not teach claim 14 is because Derakhshani captures images of the eyes and not the face. (EX2061, ¶94.) As I discussed above, Dr. Turk is wrong. I note that Dr. Turk does not rebut that Tahk teaches claim 14.

C. Derakhshani, Tanii, And Hoyos (Ground 1C)

55. Ground 1C challenges claim 18, which requires "detect[ing] a reflection of the displayed image off of the user's face."

56. Dr. Turk argues that the Derakhshani-Tanii-Hoyos combination does not teach claim 18 because Derakhshani captures images of the eyes and not the face. (EX2061, ¶¶94, 120-121.) As I discussed above, Dr. Turk is wrong.

57. Dr. Turk also argues that Hoyos detects a reflection off the user's eye, not the user's face. (EX2061, ¶122.) Dr. Turk conflates Hoyos disclosures. Hoyos discloses at least two embodiments: the [0018] embodiment and the [0019]

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

embodiment. In the [0019] embodiment, both "the first reflection [of the first image] and the second reflection [of the second image] were formed by a curved surface consistent with a human eye." (EX1010, [0019].) In the [0018] embodiment, only the "second image includes a representation of the first image on the computer screen reflected by a curved surface consistent with a human eye." (EX1010, [0018].) In other words, while both images in the [0019] embodiment are a reflection off of the user's eye, the [0018] embodiment discloses only the second image being a reflection off of the user's eye. As a result, the first image in the [0018] embodiment may be a reflection off of the user's face. Dr. Turk testified during his deposition that the reflection off the eye in the [0018] embodiment is for the second image only. (EX1059, 123:6-10.)

58. Dr. Turk's other arguments are irrelevant or incorrect. Dr. Turk's argument that the surface of an eye for reflection is not the same as the surface of a face is irrelevant because claim 18 is not limited to a specific type of surface for the reflection. (EX2061, ¶121.) And Dr. Turk's assertion that Hoyos states it detects reflections from the eye only is incorrect. (EX2061, ¶122.) Instead, Hoyos states the "system expects to see a change ***primarily on the cornea***, where a sharp image of the screen is reflected." (EX1010, [0035] (emphasis added).) Being "primarily on the cornea" means the reflection is mostly, but not entirely, on the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

cornea. Dr. Turk agreed. (EX1059, 123:11-124:8.) This is consistent with the reflection in the [0018] embodiment being off of the user's face.

D. Zhang And Tanii (Ground 2A)

1. Zhang and Tanii teach "comparing the first data to the second data to determine whether expected differences exist between the first data and the second data" ('471 Patent, Elements 1[d6] and 10[f]) and "differences between the first data and the second data have expected distortion" ('471 Patent, Elements 1[d7] and 10[g])

a. Zhang

59. Zhang teaches biometric authentication that involves distinguishing live faces from flat surfaces. (EX1006, Abstract, Title.)

60. Zhang's authentication process 200 is depicted in the flowchart of Figure 2 below. The authentication process involves accessing multiple images of the user's face (act 202) and then two parts: (1) determining whether the face is a 3D structure or a flat surface (act 204), and (2) determining whether motion is present (act 206). (EX1006, [0022], Fig. 2.) If the face is a 3D structure or motion is present than the system authenticates the user. (EX1006, [0023], Fig. 2.)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

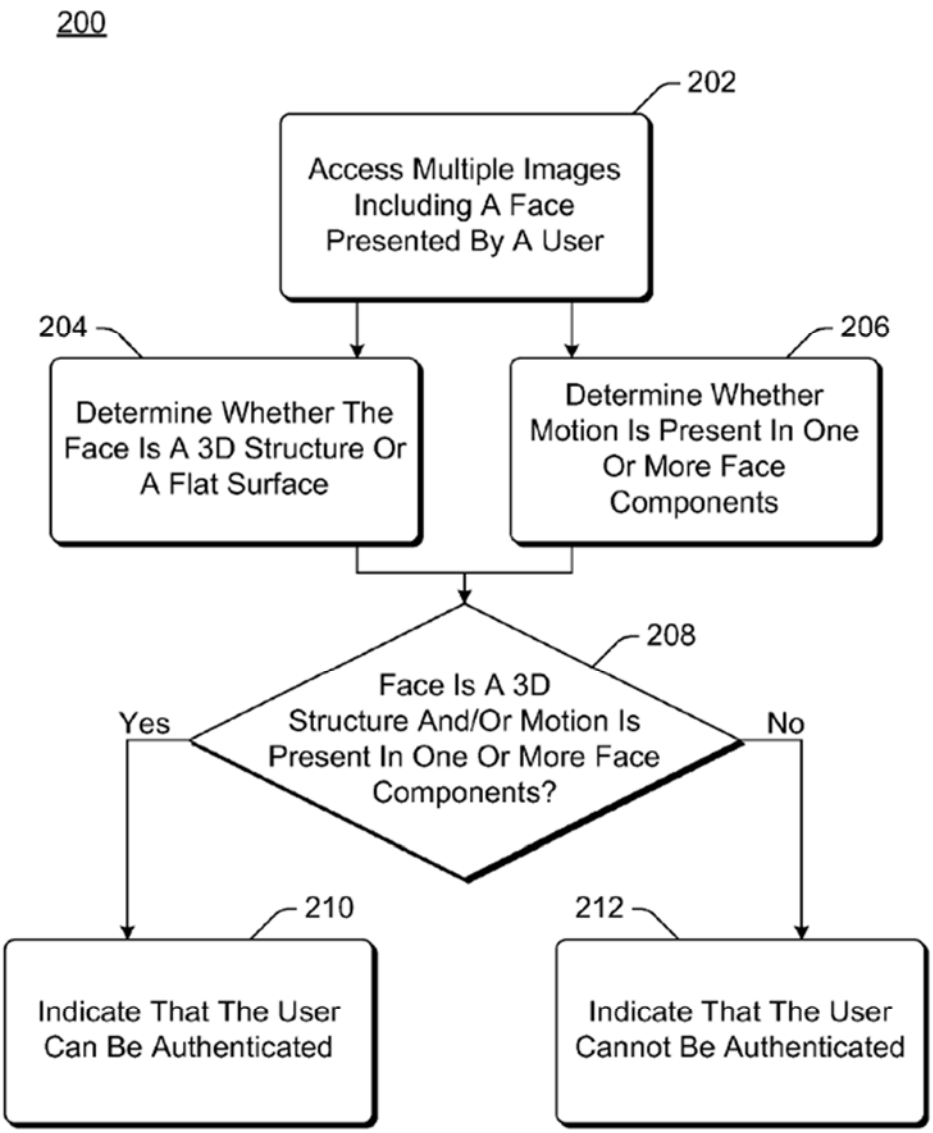


Fig. 2

61. 3D structure vs. flat surface: Process 300 shows how Zhang determines whether the face is a 3D structure or flat surface. It involves extracting feature points from two images in act 302. (EX1006, [0027].) As Zhang explains, "[t]he one or more feature points are feature points of the face in the first and

second images. A variety of different feature points can be extracted, such as a corner of an eye, a corner of a mouth, a top of a nose, and so forth." (EX1006, [0027].) The extracted feature points are matched in act 304 by identifying the locations of the same features points in the two images. (EX1006, [0028].) Then a homography matrix H is generated in act 306, and then used to generate a warped image (by warping the first image with the homography matrix H) in act 308. (EX1006, [0029]-[0030].)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

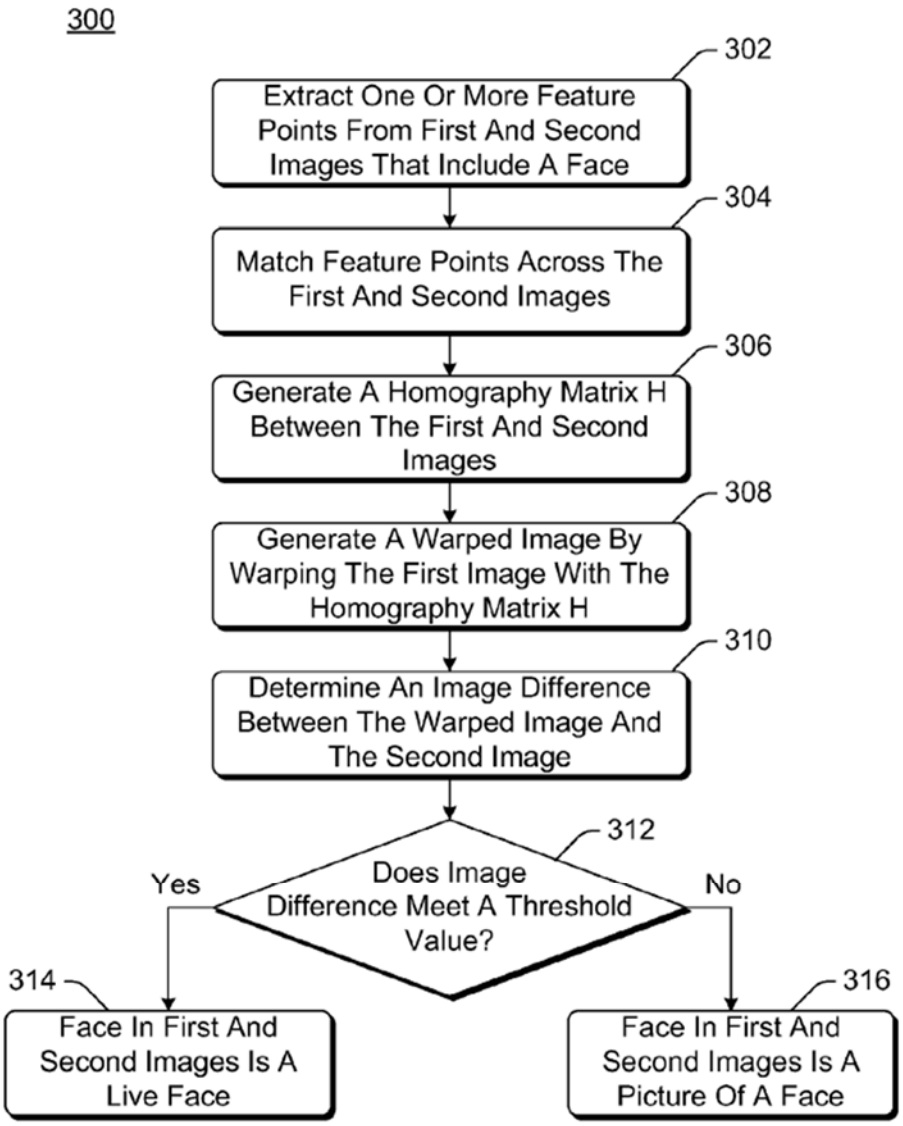


Fig. 3

62. Zhang discloses determining the difference in color/intensity values between the warped image and the second image in act 310. (Zhang (EX1006), [0031]-[0032].) If the image difference meets a threshold value (act 312), then the face is determined to be a live face (act 314). (Zhang (EX1006), [0033]-[0034].)

When the images are of a live face, the images exhibit differences in the data (such as differences in color or intensity values) between the images and also exhibit distortion as expected (expected differences and expected distortions) due to being a live face. (EX1006, [0031]-[0034], Claim 20.) These differences appear as distortion, which may be determined by looking at differences in color/intensity values of the pixels. As a result, Zhang's determination of whether the face is a live face is a determination of whether expected differences, having expected distortion, exist.

63. Motion estimation: Process 400 shows how Zhang determines whether motion is present. Process 400 includes extracting feature points (act 404), reconstructing a 3D scene (act 406), fitting the reconstructed 3D scene to a plane (act 408), fitting characteristics of the reconstructed 3D scene to a plane (act 410), determining whether characteristics meet a threshold (act 412), and determining that the face is a live face (act 414).

64. In fitting characteristics to a plane in act 410, Zhang determines whether the face is a live face by analyzing characteristics of facial components or feature points. Zhang determines a first value by determining the change in distance from a particular feature: "a maximum distance from a particular feature (e.g., tip of a nose)." (EX1006, [0045].) Zhang also determines a second value by determining the change in distance between two facial components: "the distance

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

between two facial components on the plane (e.g., the distance between the eyes or between the corners of the eyes)." (Zhang (EX1006), [0045].) Zhang calculates the ratio of the first value and the second value. (Zhang (EX1006), [0045].) A ratio that is not 1 would be due to expected distortion of a live face. A POSITA would have understood that the change in distance from a particular feature and the change in distance between two facial components are expected differences that are present when the face is a live face.

65. Dr. Turk acknowledges that Zhang teaches determining differences, but argues that it is not a determination of "expected differences" or "expected distortion." (EX2061, ¶126.) Dr. Turk reasons that color/intensity differences do not relate to expected distortion. (EX2061, ¶¶127-128.) This is false. A POSITA would have understood that color/intensity differences in images indicate depth, thereby captured images having color/intensity differences indicate a three-dimensional object such as a face exhibiting expected distortion. Zhang explains this: "[i]f the image difference meets the threshold value, then the face in the first and second images is determined to be a live face (act 314)." (EX1006, [0034].) Derakhshani explains this concept: "the intensity or color scale can be zoomed in to reveal more subtle changes in the level of the detected image signal. If the captured images are of a ***live eye, it is expected that the range of color or intensity values detected will continue to vary continuously***. In contrast, a spoofed image

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

(e.g., a digital photograph presented to the sensor) may exhibit large discontinuous jumps corresponding to halftones."). (EX1005, 19:30-37.) Additionally, Dr. Turk's focus on a single pixel in each image is irrelevant because Zhang still determines the differences.

66. I further disagree with Dr. Turk's arguments that the color/intensity differences identified in Zhang does not relate to expected distortion that is based on movement of the camera or the user, and that Zhang makes no prediction of how an image would change due to this movement. (EX2061, ¶127.) This argument mischaracterizes Zhang's teachings and the Petition's Zhang-Tanii combination. Zhang discloses accessing "multiple images including a face presented by a user" and determining whether the face is a "3-dimensional structure or a flat surface" by analyzing differences between images. (EX1006, Abstract, [0025].) Zhang explains that when two images of a flat surface are compared, they can be aligned using a homography matrix such that "the warped image will be similar to the second image" with a "difference between the second image and the warped image [that] is small." (EX1006, [0025].) In contrast, when the images depict a three-dimensional face, "the difference between the second image and the warped image is larger" because the homography matrix cannot align different views of a 3D object. (EX1006, [0025].) The color/intensity differences that Zhang measures are the manifestation of this misalignment. A

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

POSITA would have understood that these color/intensity differences arise because different parts of the three-dimensional face are at different distances from the camera, causing them to appear differently across viewpoints in ways that a planar transformation cannot correct.

67. Moreover, the Zhang-Tanii combination does not rely on Zhang alone to teach analyzing expected distortion resulting from camera or user movement. Rather, as I explained in my opening declaration, the Zhang-Tanii combination relies on Tanii for teaching that three-dimensional faces exhibit predictable, distance-dependent perspective distortion. (EX1003, ¶¶260-270.) Tanii discloses "where the main object 9 and the cellular phone 1 are close together … an unnatural image results in which the perspective is exaggerated," and explains the technical reason: "the difference between the angle θ1 … and the angle θ2 … decreases as the lens unit 211 comes closer to the main object 9." (EX1007, [0047]-[0049].) Tanii further discloses that "the warp or distortion increases as the angle θ1 … increases" and that "where the main object is close to the imaging portion 2 and the image of the main object occupies a large percentage of the entire image, the exaggeration warp of the image becomes substantial." (Tanii, [0049].) These teachings provide the prediction Dr. Turk claims is missing: as the camera moves closer to a three-dimensional face, the perspective distortion increases in a predictable manner, with "peripheral areas of the main object appear[ing] reduced

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

in size relative to the center area." (EX1007, [0056].) A POSITA would have understood that Zhang's framework would be enhanced by Tanii's teachings that the type and magnitude of differences expected from a three-dimensional face vary predictably with distance due to perspective distortion. The combination teaches comparing observed differences (as Zhang does via color/intensity analysis) to expected differences (as Tanii characterizes via distance-induced distortion) to determine three-dimensionality for authentication. Dr. Turk's argument that the color/intensity differences do not relate to expected distortion based on movement fails to account for Tanii's contributions to the combination.

68. Dr. Turk alleges that Jumio inconsistently maps Zhang's "feature points" (elements 1[d2] and 10[b]) and "facial-feature location" (elements 1[d7] and 10[f]) to the "first data." (EX2061, ¶¶129-134, p. 67 subheading.) As I explained in my opening declaration, the extracted feature points are data that characterize an individual's unique physical characteristics. (EX1003, ¶¶289-290.) The feature point location characterizes the location of the individual's unique physical characteristics. In other words, Zhang's feature point location is part of the feature point data, thereby showing that the mapping is not inconsistent.

69. And Dr. Turk's arguments that the Petition cites to different passages in Zhang is flawed. (EX2061, p. 70 fn. 2.) Elements 1[d2] and 10[b], and elements 1[d7] and 10[f] recite different limitations. Elements 1[d2] and 10[b] are

directed to processing to create first data, while elements 1[d7] and 10[f] are directed to differences between first data and second data. Both sets of passages, [0027] (relied on for elements 1[d2] and 10[b]) and [0031]-[0032] (relied on for elements 1[d7] and 10[f]), discuss feature point data.

70. Further, Dr. Turk incorrectly tries to differentiate between feature point data used to generate the homography matrix H and feature point data used in the comparison of images. (EX2061, ¶¶131-135.) However, as Zhang explains, "feature points" are parts of the face, "such as a corner of an eye, a corner of a mouth, a tip of a nose, and so forth" that are extracted from the first and second images. (EX1006, [0027].) The extracted feature points are matched across the first and second images and used to generate the homography matrix H. (EX1006, [0028]-[0029].) Zhang discloses "[t]his matching of the feature points across the first and second images refers to identifying the locations of the same feature points in each of the two images. For example, the locations of the same corner of the same eye in the face in the first and second images are identified." (EX1006, [0028].) Then, when comparing the images, the image difference computes the "difference in color or intensity values at the same pixels (locations) in the two images." (EX1006, [0031].) Zhang uses the same term "locations" when describing the generation of the homography matrix H and also when describing the difference in color/intensity values because it is referring to the same extracted

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

feature points. Using Zhang's example of the corner of the eye as the "location," it would make sense for the comparison between the two images to be a comparison of the same feature point, such as the same corner of the same eye because otherwise a difference in color/intensity value may be due to different features points having different colors or intensity values due to, e.g., the user moving. For example, the color of the corner of an eye is likely to be different than the color of the corner of a mouth, thereby a comparison of the same pixel position having two different feature points (as Dr. Turk suggests) would not be indicative of whether the face is a 3D structure or a flat surface. Thus, Zhang's feature points are used for both generating the homography matrix H and the comparison of the images (part of authenticating the user).

71. Dr. Turk alleges that a three-dimensional image would not align with Zhang's approach and use of a homography matrix, and then concludes that Zhang does not predict how a three-dimensional face would distort across different views. (EX2061, ¶68.) However, Zhang's homography-based approach recognizes the fact that a three-dimensional face produces characteristic differences across views that are not eliminated by planar transformations. Zhang teaches that "two views of a flat (planar) surface are related based on a homography matrix H" such that warping one image using the homography matrix H produces an image similar to the second image. (EX1006, [0025].) If there is a large difference between the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

first and second images, Zhang determines that "the first and second images include a face that is a live face rather than a flat surface." (EX1006, [0025].) A large difference is due to the images being of a three-dimensional face because different parts of the face are at different depths. And a small difference indicates a flat surface (expected for a 2D photograph). (EX1006, [0025], [0033]-[0034].) A POSITA would have understood that Zhang's approach involves three-dimensional objects (including faces) that exhibit view-dependent distortions that are distinguishable from two-dimensional objects.

72. Moreover, as I explained above and in my opening declaration, the Petition combines Zhang with Tanii's teaching of the distance-induced distortion of a face when the camera comes closer to the face. (EX1003, ¶¶260-270.) Tanii teaches that perspective distortion of three-dimensional faces varies predictably with camera-to-subject distance, with closer distances producing more pronounced "exaggeration warp" where "peripheral areas of the main object appear reduced in size relative to the center area." (EX1007, [0047]-[0049], [0056].) The Zhang-Tanii combination teaches a POSITA that (1) three-dimensional faces produce characteristic differences across images captured at different distances (Zhang's principle) and (2) these differences manifest as distance-induced perspective distortion (Tanii's teaching). Together, the Zhang-Tanii combination teaches

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

determining whether expected distortion exists in the captured images, which is what the ’471 claims require.

b. Tanii

73. Dr. Turk makes arguments about Tanii for the Zhang-Tanii combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii combination. (EX2061, ¶¶136-137, 149.) I explained why Dr. Turk’s arguments are incorrect above.

2. A POSITA would have been motivated to combine Zhang and Tanii

74. Dr. Turk argues that there is no motivation to combine Zhang and Tanii because Tanii destroys Zhang’s “basic objective” that requires distortions to operate. (EX2061, ¶¶162-163, 166.) Dr. Turk’s arguments are however premised on his misunderstanding of the Zhang-Tanii combination. The teachings of Tanii that are part of the Zhang-Tanii combination are its teachings of distance-induced distortions in images of a user’s face. (EX1003, ¶¶260-270.) Dr. Turk erroneously bases his arguments on Tanii’s teachings of image correction, which are not part of the Zhang-Tanii combination. (EX2061, ¶¶165, 173-174, 178.)

75. Dr. Turk also criticizes Tanii’s figures for allegedly being inconsistent with its disclosures. (EX2061, ¶176.) Specifically, Dr. Turk alleges that the ears in the warped image (Figure 9) would not “miraculously appear after Tanii’s purported distortion correction.” (EX2061, ¶176.) However, Dr. Turk’s argument

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

shows that he misunderstands Tanii. A POSITA would have understood that Tanii's figures are simplified line drawings that portray abstract representations of the distance-induced distortions caused by the optical effects. (EX1007, [0048], [0049], [0056].) An actual captured image, on the other hand, includes far more information. Thus, Tanii's figures do not depict the level of detail of a captured image. Tanii does not cause ears to "miraculously appear," as Dr. Turk asserts. Instead, barrel (aka perspective) distortion is portrayed: the magnification is changing based on the distance from the image center due to how the camera lens works. (EX1003, ¶¶46-51.) Given that Tanii's figures are simplified and abstract depictions, a POSITA would have understood that the ears are present in the actual captured image corresponding to Figure 4B, but because of the magnification, the ears are difficult to see. For example, the rightmost image (below) has ears that are present, but is difficult to see with the human eye due the distance-induced distortion. In those photographs, the user's ears appear increasingly distorted (reduced in lateral size compared with other facial features) as the distance between the sensor and the user is reduced. Also consistent with Tanii's figures, the user's nose exhibits distortion where it is relatively enlarged as this distance is reduced.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

3. Zhang and Tanii teach "interpolating the first data and the second data to obtain estimated intermediate data" ('471 Patent, Elements 2[a] and 11[a])

76. Dr. Turk argues that Zhang does not teach elements 2[a] and 11[a] because it does not model expected distortion. (EX2061, ¶139.) However, the '471 claims do not require modeling expected distortion. Elements 2[a] and 11[a] require interpolating first and second data to obtain estimated intermediate data, which Zhang teaches. As I explained in my opening declaration, Zhang teaches additional intermediate images existing between the two images being compared for authentication. (EX1003, ¶¶309-313; Zhang, [0035].) The intermediate images are used for feature point extraction and feature point matching. (EX1006, [0036].) A POSITA would have been motivated to generate predictions of what the intermediate images should look like, thereby interpolating the data of the two images to obtain estimated intermediate data.

4. Zhang and Tanii teach "verifying the presence of the user's ears in the at least one first image, and verifying the absence or reduced visibility of the user's ears in the at least one second image" ('471 Patent, Claims 3 And 12)

77. Dr. Turk makes arguments about Tanii for the Zhang-Tanii combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii combination. (EX2061, ¶¶140-141.) As I explained above, Tanii teaches claims 3 and 12.

5. Zhang and Tanii teach "the first data and the second data comprise biometric data" ('471 Patent, Claims 7 and 16)

78. Dr. Turk argues that Zhang does not teach claims 7 and 16 because its feature points are not "biometric data" used for authentication. (EX2061, ¶¶142-144.) According to Dr. Turk, Zhang's feature points are used only to generate a homography matrix H. (EX2061, ¶142.) This is the same argument that Dr. Turk made for elements 1[d6] and 10[f], and elements 1[d7] and 10[g]. I explained why Dr. Turk's arguments are incorrect above. Further, Zhang's facial features meet the description in the '471 specification of "biometric data," which Dr. Turk agrees with. (EX1059, 50:11-14.) Zhang states that "feature points are feature points of the face," "such as a corner of an eye, a corner of a mouth, a top of a nose, and so forth." (EX1006, [0027].) Thus, the comparison involves biometric data.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

E. Zhang, Tanii, And Hoyos (Ground 2C)

79. Ground 2C challenges claim 18, which requires "detect[ing] a reflection of the displayed image off of the user's face." Dr. Turk makes arguments about Hoyos for the Zhang-Tanii-Hoyos combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii-Hoyos combination. (EX2061, ¶¶145-148.) As I explained above, Hoyos teaches claim 18.

F. Tanii Is Analogous Art

1. Tanii is from the same field of endeavor

80. Dr. Turk argues that Tanii is not in the same field of endeavor as the '471 patent because the '471 uses distortion and Tanii tries to remove it as part of its image correction process. (EX2061, ¶¶150-151.) But Dr. Turk's argument that Tanii assumes distortion exists and corrects for it is misplaced. The Derakhshani-Tanii combination and the Zhang-Tanii combination rely on Tanii for its teachings of distance-induced distortion, not on its teachings of an image correction process.

81. Further, a POSITA would not have understood Tanii's teachings to be so narrow as Dr. Turk argues. Instead, a POSITA would have understood that one of Tanii's teachings is related to identifying and accounting for distance-induced distortions of three-dimensional faces in images captured using a camera.

82. Thus, both the '471 patent claims and Tanii relate to distance-induced distortion. That is, Tanii is directed to the same function (accounting of distance-

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

induced distortions) and structure (camera), and as a result is in the same field of endeavor as the ’471 patent.

2. Tanii is reasonably pertinent to the problems faced by the inventors of the ’471 Patent

83. Dr. Turk acknowledges that one of the problems the ’471 inventors were faced with is “this long-understood issue” of a three-dimensional face exhibiting distortions when the camera and face are close together during image capture. (EX2061, ¶156.) Dr. Turk does not dispute that Tanii is reasonably pertinent to this distance-induced distortion problem, but instead argues that Tanii’s teachings would not be helpful because the problem was known. (EX2061, ¶156.) But whether Tanii’s teachings were known does not determine whether Tanii is reasonably pertinent.

84. I note that the Petition stated that “[t]he ’471 Patent describes . . . methods to distinguish a three-dimensional face from a two-dimensional picture of a face.” (Petition, 4.) The Petition did not admit that distinguishing a three-dimensional face from a two-dimensional picture of a face was a problem faced by the ’471 patent, as Dr. Turk alleges. (EX2061, ¶155.)

85. Dr. Turk argues that the ’471 patent solves the spoofing problem and Tanii does not. (EX2061, ¶¶157-158.) But spoofing is not the only alleged problem faced by the ’471 inventors. As discussed above, Tanii solves the ’471 problem related to distance-induced distortions.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

86. Dr. Turk criticizes Tanii because allegedly "the '471 patent uses the distortion that Tanii tries to remove." (EX2061, ¶151.) He argues that "Tanii works against the challenged claims, by removing the very thing that enabled the patent's inventor to solve the spoofing problem." (EX2061, ¶160.) I disagree. A POSITA would have understood that Tanii's teachings of distance-induced distortion is suitable for use for user authentication.

87. And Dr. Turk's argument about purpose is flawed. (EX2061, ¶161.) As described in the context of the relevant art is discussed, [0006] states "[i]f the image of the face of a person or the like is captured from a point that is relatively nearby, an image is obtained in which the perspective appears exaggerated. . . . When a camera comes close to a three-dimensional object that has depth, the peripheral areas of the object, which are farther away from the camera than the front area, appear to be closer to the front area than they really are, resulting in a warp that causes the image to appear as if the perspective were exaggerated (the type of warp in which the perspective becomes exaggerated, as in this case, will hereinafter be termed 'exaggeration warp', as distinguished from distortion)." (EX1007, [0006].) Thus, Tanii accounts for warp that causes the image to appear as exaggerated.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

IV. ’606 PATENT

A. Derakhshani And Tanii (Ground 1A)

1. Derakhshani and Tanii teach “verifying three-dimensionality of a user’s face using images of the user’s face” (’606 Patent, Elements 1[pre], 10[pre], and 19[pre])

a. Derakhshani

88. As I explained in ¶¶9-16, Derakhshani teaches biometric authentication that involves determining whether images depict a three-dimensional subject (user).

89. Dr. Turk agreed that Derakhshani teaches authenticating three-dimensionality of a user. (EX1059, 128:21-129:1.) He argues that Derakhshani does not teach capturing an image of the user’s face. (EX2061, ¶¶179, 185.) I disagree.

90. Dr. Turk does not dispute that Derakhshani teaches capturing at least part of a user’s face. He however argues that capturing part of the user’s face (other than the eye) is incidental. (EX2061, ¶194.) I do not agree that capturing part of the user’s face is incidental.

91. Derakhshani discloses determining a spatial metric “based on a distance from a sensor to a landmark that appears in a plurality of the images each having a different respective focus distance.” (EX1005, 16:44-47.) The determined spatial metric is used to determine three-dimensionality: “[a] high

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

value for this spatial metric may indicate a three-dimensional subject and a higher likelihood that the subject is a live eye, while a low value may indicate a higher likelihood that the subject is a two-dimensional spoof attack." (EX1005, 17:22-26.)

92. Derakhshani specifically names "a nose, an ear," and a "mouth" as example landmarks identified in the captured images. (EX1005, 16:51-54, 17:16-17.) The passage at 17:29-36 discusses "landmarks on a particular user's face." (EX1005, 17:29-36.) And the passage at 3:13-15 states that the "landmark may be a portion of a face depicted in the images." (EX1005, 3:13-15.) Derakhshani also states that an image may comprise multiple "landmarks (e.g., an eye, nose, mouth, and ear." (EX1005, 17:16-19; *see also* 3:5-7 ("two or more landmarks that appear in a plurality of the images"), 17:2-6 ("multiple landmarks").) Dr. Turk also agreed that Derakhshani captures an image having multiple landmarks. (EX1059, 136:25-137:3.) It would have been obvious for an image comprising an eye, nose, mouth, and ear, as Derakhshani discloses, to be an image of a face because these are facial features that when they are all included in an image, it would be an image of a user's face. (EX1005, 17:16-20.)

93. Derakhshani also names a "background object" as an example landmark. (EX1005, 16:51-54.) It would have been obvious for an image comprising an eye and a background object (landmark) to be an image of a user's

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

face. For example, the background object may be a painting on a wall located behind the user. In order to capture the user's eye and the painting in the image, the sensor would have to be back far enough from the user such that the user's face would be captured. Thus, Dr. Turk's suggestion (at, *e.g.*, ¶¶187 and 191 of his declaration (EX2061)) that Derakhshani captures images of the user's eyes only is wrong. I also note that Derakhshani does not disclose capturing the user's eye vasculature as its primary purpose, as Dr. Turk argues at ¶¶189 and 194 of his declaration.

94. Derakhshani further states that its authentication is "based ***in part*** on one or more images of the white of an eye 410." (EX1005, 5:65-67.) Being based "in part" means that some of the captured images are of the white of an eye 410 (such as the images captured using the sensor 6-12 cm from the eye (EX1005, 13:35-40)), and some of the captured images are not just of the white of an eye, such as images of the user's face.

95. Dr. Turk's declaration argument that the spatial metric determination relies on characteristics of the user's eye only is also incorrect. (EX2061, ¶188 (citing EX1005, 4:19-21).) Derakhshani's spatial metric involves capturing a plurality of images at difference distances. (EX1005, 16:44-47, 17:45-47.) One image of the plurality of captured images may be a close-up image, and another image of the plurality of captured images may be captured at a different distance

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

further away from the user. The close-up image may capture the eye vasculature (with enough precision), and the further-away image may capture the landmark(s) (with enough precision).

96. The passage at 4:19-21 that Dr. Turk cites describes image(s) used for determining the match score. The match score is determined using image(s) that capture the vasculature of the eye and are "subjected to further analysis" (step 616) ***after and separate from*** the determination of the spatial metric (step 714, which is part of step 710 performed as step 604). (EX1005, 14:59-64, 16:44-18:7.) As Dr. Turk testified, the eye vasculature is determined after determining the liveness score and determining the liveness score is separate from determining the match score. (EX1059, 130:3-7, 132:9-11.) Further, Dr. Turk agreed during his deposition that the spatial metric determination is not solely based on characteristics of the user's eye. (EX1059, 126:4-13.) Dr. Turk also agreed that Derakhshani captures images having more than the user's eye. (EX1059, 136:16-24.)

97. Further, a POSITA would have understood a prior art reference based on not only what is depicted in the figures, but also the teachings discussed in the specification. Even Dr. Turk looks to the specification when reviewing a prior art reference. (EX1059, 16:3-17:24.) Accordingly, Dr. Turk's declaration arguments that Derakhshani captures images of the eyes only because that is what is depicted

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

in the figures is without merit because Derakhshani's specification suggests otherwise. (EX2061, ¶¶179, 189-190; EX1005, 16:44-54, 17:16-17.) Dr. Turk also testified that if there is an embodiment with the image including a nose, ear, mouth, and eye, it would be an image that is not depicted in Derakhshani figures. (EX1059, 137:4-17.) Indeed, Derakhshani discloses such an embodiment, and thus, its figures do not show all of the disclosed embodiments. (EX1005, 17:16-19.)

98. Dr. Turk argues that Derakhshani's system is not adapted to capture the user's face or head. (EX2061, ¶¶189, 191.) I disagree. To make an accurate determination of the spatial metric, a POSITA would have understood that there must be a discernable difference in the degree of focus (for the focus-distance approach) or movement of the landmarks (for the parallax approach), which would mean there must be a discernable difference in distance between the user and the sensor. This would mean at least one captured image would be taken such that the sensor is far enough away from the user, thereby including more than just the user's eye in the image. Derakhshani discusses this at 17:55-59: "[i]f all the landmarks in the image undergo the ***same apparent displacement*** due to relative motion of the sensor, ***i.e., the difference in the parallax effect for the landmarks is small***, then the subject viewed by the camera has [a] higher likelihood of being a two-dimensional spoof attack." (EX1005, 17:55-59.) A POSITA would have

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

understood from this passage that a large difference in the parallax effect for the landmarks would be needed when the object is three-dimensional in order to avoid a false determination of a two-dimensional spoof attack.

99. Further, Derakhshani teaches achieving a finer resolution by using a reduced dynamic range for the sensor. "[O]ne or more images may be captured using a reduced dynamic range for the sensor (e.g., a camera) so that a finer resolution in intensity of detected light is achieved in a range within which it occurs in the captured images. In this manner, the intensity or color scale can be zoomed in to reveal more subtle changes in the level of the detected image signal." (EX1005, 19:24-32.) Thereby, by adjusting the parameters, the same sensor may be used to capture images of the vasculature of an eye with a finer resolution and images of the face. This means Derakhshani's system is adapted to capture the user's head/face.

100. Dr. Turk also argues that simultaneous capture of the eye vasculature and face in the same image with enough precision for both is required. (EX2061, ¶193.) I disagree based on Derakhshani's disclosures of capturing a plurality of images for the spatial metric determination. For the focus-distance approach, the plurality of images have different respective focus distances for a given landmark. Derakhshani provides an example where the landmarks in an image include "an eye, nose, mouth, and ear." (EX1005, 17:16-17.) It explains that a landmark may

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

be "in focus" or may be blurred (also described as "the degree of focus is low"). (EX1005, 16:56-66.) A POSITA would have understood Derakhshani's disclosures of different focus distances in different images and a landmark being in focus and blurry means the landmark is in focus in one image and blurred in another image. The spatial metric is determined based on distances of the landmarks. (EX1005, 16:66-17:6, 17:12-22, 17:49-52.) A high value for the spatial metric indicates a three-dimensional subject and a low value indicates a two-dimensional spoof attack. (EX1005, 17:22-26.) A POSITA would have understood that large estimated distances from the landmarks would occur when some of the landmarks are blurred (out of focus) and result in a high value for the spatial metric. Thus, Derakhshani does not need to capture images that simultaneously capture both the vasculature of an eye and a face with enough precision in order to determine three-dimensionality or authentication, as Dr. Turk argues. (EX2061, ¶193.)

101. Instead, a POSITA would have understood that Derakhshani may be able to determine three-dimensionality when not simultaneously capturing both the vasculature of an eye and face with enough precision. For example, Derakhshani may capture a close-up image with the user's eye in focus but the ear is not in focus, and may also capture another image that is further away from the user where the ear is in focus. The spatial metric may be determined using the close-up image

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

and the further-away image, both having a discernable difference in distance of the ear landmark, thereby resulting in a high spatial metric and determination of three-dimensionality. The match score may be determined using the close-up image. In such a situation, an image that simultaneously captures both the eye vasculature and the face with enough precision is not required. Thus, not every image needs to be close enough to capture the vasculature of the eye with sufficient clarity, as Dr. Turk argues. (EX2061, ¶¶180, 184.)

a. Tanii

102. Tanii also teaches capturing images of a user's three-dimensional face involving distance-induced distortions, as shown in Figures 3A-3B and 4A-4B (reproduced below). (EX1007, Abstract ("an image is captured with the user's face at a close range"), [0006], [0013], [0047]-[0049], [0061]-[0063].) Dr. Turk does not dispute this, but rather argues that Tanii does not identify a face. (EX2061, ¶200.) This is incorrect. Tanii states that it "detects the size of the image of the main object." (EX1007, [0074].) Tanii explains that "the image of the main object, particularly the image of the face of the user who is holding the cellular phone 1, is captured." (EX1007, [0061].) Thus, Tanii detects the size of the user's face. In order to detect the size of the user's face, it must first detect the user's face. Further, Tanii describes different ways to detect the user's face, such as "extracting the contours of the main object using image processing." (EX1007,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

[0069].) Tanii also describes "identif[ying] the region of the image of the main object based on the location of clear edges in the image as well as on the color distribution in the image." (EX1007, [0076].)

Fig.3A

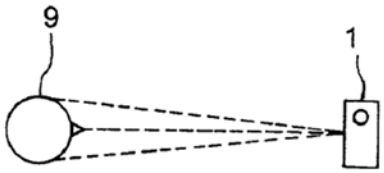


Fig.3B

Fig.4A

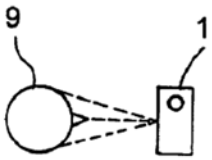
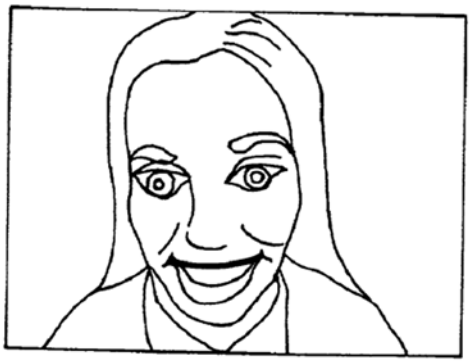


Fig.4B



2. Derakhshani and Tanii teach "biometric data" ('606 Patent, Elements 1[b], 10[c], and 19[a])

103. Dr. Turk argues that Derakhshani does not teach elements 1[b], 10[c], and 19[a] of the '606 patent because its landmarks are not "biometric data." First,

I disagree with Dr. Turk's separation of landmarks into what he calls "image landmarks" and "anatomical landmarks." (EX2061, ¶58.) Dr. Turk seems to create this separation because "image landmarks" are used for the liveness score determination and "anatomical landmarks" are used for the match score determination. (EX2061, ¶58.) However, Derakhshani teaches the eye as being an example landmark used for both the liveness score determination and the match score determination, thereby showing that a landmark cannot be classified as either an "image landmark" or an "anatomical landmark." (EX1005, 14:25-29; 17:16-19.)

104. Dr. Turk argues that Derakhshani's landmark is not "biometric data" because it may be an iris, eye corner, or ear, which is not used to identify a user. (EX2061, ¶199.) But Derakhshani uses the iris, eye corner, or ear to determine a liveness score or a match score, both of which are used to authenticate the user. (EX1005, 11:46-51, 16:44-54, 17:12-26, 19:16-18, 21:8-28, Figs. 6-8.)

105. I note that Dr. Turk's requirement provided in his declaration that "biometric data" be used to identify a user is inconsistent with the '606 specification, which states: "[t]he term biometric data may include among other information biometric information concerning facial features (distorted or undistorted) and ***path parameters***. ***Examples of path parameters may include an acceleration and speed of the mobile device, angle of the mobile device during***

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

image capture, distance of the mobile device to the user, path direction in relation to the user's face position in relation to the user, or any other type parameter associated with movement of the mobile device or the user face in relation to a camera. Other data may also be included such as GPS data, device identification information, and the like." ('606 patent (EX1001), 6:54-64.) This is inconsistent with his testimony where he agreed with the specification's description of "biometric data." (EX1059, 50:11-14.)

3. Derakhshani and Tanii teach "comparing[/compare] the first biometric data to[/with] second biometric data to determine whether differences between the at least one first image and the at least one second image match expected differences" ('606 Patent, Elements 1[f], 10[g], and 19[d])

106. Tanii teaches the "expected differences" of elements 1[f], 10[g], and 19[d]. (EX1007, Figs. 3A-3B, 4A-4B, 5, [0047]-[0049].) Tanii explains that when there is a sufficient distance between the main object 9 and the cellular phone 1, the image appears natural. (EX1007, [0047], Figs. 3A-3B.) When the main object 9 and cellular phone 1 are too close to each other, the perspective is exaggerated resulting in an unnatural image. (EX1007, [0047], Figs. 4A-4B.) This exaggeration in perspective and unnaturalness is an expected distortion because it occurs due to known and expected differences that result from the light ray striking different points of the main object 9. (EX1007, [0048]-[0049], Fig. 5.)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

107. Dr. Turk does not dispute these teachings. Instead, he focuses his arguments on whether Tanii evaluates degree of distance-induced distortion, or measures, assesses, or undoes three-dimensional distortion. (EX2061, ¶¶182, 202.) But the '606 claims do not require evaluating the degree of distance-induced distortion, or measuring or undoing three-dimensional distortion. Further, the Derakhshani-Tanii combination relies on Derakhshani for its teachings of authenticating the user, which Dr. Turk does not dispute Derakhshani teaches.

108. Moreover, Dr. Turk is incorrect that Tanii does not assess three-dimensional distortion and that it assumes an image requires correction. (EX2061, ¶246.) Tanii determines the distance from the main object to the camera. (EX1007, [0101].) It discloses "where the distance to the main object equals or exceeds the threshold value D1, it is determined that the main object and the imaging portion 2 are located a sufficient distance apart and that as a result no correction is needed." (EX1007, [0101].) That is, when there is no distance-induced distortion, Tanii does not perform a correction. Tanii also teaches determining the degree of distance-induced distortion and applying the appropriate amount of correction. It discloses "[w]here the distance to the main object is less than the threshold value D1" or "less than the threshold value D2," then "a low degree of correction" or "a high degree of correction" is needed. (EX1007, [0101].) It further discloses "[i]t can be automatically determined whether or not

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

correction is needed." (EX1007, [0136], Claim 7.) In other words, Tanii identifies and assesses distance-induced distortions.

109. Dr. Turk also argues that Derakhshani compares the captured images to an expected plane/shape, but does not compare data of two captured images together. (EX2061, ¶186.) This is incorrect. Dr. Turk points to 3:1-6 and 17:27-36 to support his argument, but these passages describe one embodiment. There are other passages in Derakhshani, such as 16:51-54 and 17:45-47, that teach comparing data of two captured images together. The 16:51-54 passage and surrounding text describe the spatial metric being determined based on "a distance from a sensor to a landmark that ***appears in a plurality of the images*** each having a different respective focus distance." (EX1005, 16:44-54 (emphasis added).) It states that "the focus distance may be adjusted for different images." (*Id.*; *see also* 16:44-54 ("the plurality of the images with different focus distances").) The 17:45-47 passage and surrounding text describe "parallax of two or more landmarks that ***appear in a plurality of the images***," where the "plurality of images [are] taken from different perspectives on the subject" and "the sensor is moved about the subject during image capture." (EX1005, 17:45-64 (emphasis added).) Thus, Derakhshani captures a plurality of images having different distances, and the data from the images are compared to each other.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

110. Dr. Turk further argues that Tanii aesthetically corrects images and states that its alleged "process of enlargement" does not make up for the deficiencies of Derakhshani and is incompatible with Derakhshani. (EX2061, ¶¶200, 202.) However, Dr. Turk misunderstands the Derakhshani-Tanii combination which does not use Tanii's aesthetic/warp correction. The Derakhshani-Tanii combination includes Tanii's teachings of distance-induced distortions in images of a user's face. (EX1003, ¶¶379-388.) Further, whether Tanii operates the same way when the image does not include a face is irrelevant because Tanii teaches capturing and detecting the user's face. (EX2061, ¶201; EX1007, [0062].)

4. A POSITA would have been motivated to combine Derakhshani and Tanii

111. Dr. Turk argues that there is no motivation to combine Derakhshani and Tanii because Tanii destroys Derakhshani's "basic objective" that requires distortions to operate. (EX2061, ¶¶243-244, 247.) Dr. Turk's arguments are however premised on his misunderstanding of the Derakhshani-Tanii combination and the Petition's motivation to combine. The teachings of Tanii that are part of the Derakhshani-Tanii combination are its teachings of distance-induced distortions in images of a user's face—not teachings of image correction. (EX1003, ¶¶379-388; EX2061, ¶246.)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

112. Additionally, Dr. Turk's argument that images at different distances having different distortion is not a motivation to combine again shows that he misunderstands the Derakhshani-Tanii combination. (EX2061, ¶¶203, 248.) As I explained in my opening declaration, Tanii teaches an obvious alternative to Derakhshani's focus-distance and parallax approaches. (EX1003, ¶¶379, 383.) Derakhshani's approaches require a camera with extreme sensitivity, not suitable for mobile phones that have wide angle lenses to capture a wide field of view. (EX1003, ¶384.) I note that Dr. Turk acknowledges that Derakhshani's approach has limited applicability. (EX2061, ¶248.) Dr. Turk also acknowledges that Tanii's teachings apply in situations where images are captured using a wide-angle lens. (EX2061, ¶249.) Thus, a POSITA would have been motivated to use Tanii's method when using a mobile device that has a wide-angle lens to capture the user's image(s). (EX1003, ¶¶385-388.) Dr. Turk fails to rebut this motivation.

113. I disagree with Dr. Turk's argument that Derakhshani and Tanii are incompatible. (EX2061, ¶249.) Dr. Turk's argues that this alleged incompatibility is due to Derakhshani and Tanii having different imaging requirements. (EX2061, ¶249.) Dr. Turk argues that Derakhshani needs to capture the eye vasculature with sufficient clarity and the camera must be positioned very close to the user's eye, eliminating the ability to capture the user's face or the ability to use Tanii's

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

correction method designed for capturing images at an arms length. This argument misunderstands Derakhshani's disclosures and the Derakhshani-Tanii combination. Derakhshani's system captures a plurality of images at different conditions. As I explained above, in Derakhshani, different images may be used by different functions of its authentication process. For the spatial metric determination (part of the liveness detection), Derakhshani uses images captured at different focus distances, or from different perspectives to analyze landmarks such as "a nose, an ear," "a mouth," or a "background object." (EX1005, 16:51-54, 17:16-19.) A POSITA would have understood that capturing an image containing facial landmarks results in capturing an image of the user's face. For the match score determination (eye vasculature comparison), Derakhshani uses close-up images to resolve the eye vasculature with sufficient precision. (EX1005, 13:35-40.) A POSITA would have understood that Derakhshani's multi-image framework accommodates both types of captured mages: some images optimized for the spatial metric determination (which may include facial features at varying distances from the sensor), and some images optimized for vasculature matching (which may include images of the eye vasculature). Dr. Turk assumes that all images must simultaneously capture both the eye vasculature and the facial landmarks with precision, but Derakhshani does not disclose or suggest such a requirement. Further, Dr. Turk's argument about incompatibility supports the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

motivation to combine because Derakhshani's approach is limited and applying Tanii's teachings would result in improving the implementation of biometric authentication in a mobile device as Derakhshani already envisions. (EX1005, 5:23-26; EX1003, ¶¶383-388.) Derakhshani's focus-distance and parallax approaches use specialized camera capabilities that may not be suitable for mobile phones with wide-angle lenses. Tanii teaches capturing images of the user's face using a cellular phone with a wide-angle lens at an arms length. (EX1007, [0007], [0047].) Tanii discloses that when "the user's face is close to the camera" with "a wide-angle focal length," "the exaggeration warp becomes more noticeable." (EX1007, [0007].) As I explained my opening declaration, a POSITA would have wanted to implement Derakhshani's authentication process on a mobile device, and would have been motivated to use Tanii's teachings about distance-induced distortion of an image of a user's face as an alternative or supplemental way to determining spatial metric. Derakhshani teaches using "a reduced dynamic range for the sensor . . . so that a finer resolution in intensity of detected light is achieved," demonstrating that the same sensor can be adjusted to capture different types of images by modifying camera parameters. (EX1005, 19:24-32.) A POSITA would have understood that similar adjustments would allow Derakhshani's system to capture both close-up images for the match score determination and further away images for the spatial metric determination (using

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Derakhshani's and/or Tanii's approach). As I explained in my opening declaration, a POSITA would have known how to apply Tanii's teachings to Derakhshani's system using well-known optics principles relating to camera design, and there would have been a reasonable expectation of success. (EX1003, ¶386.) Dr. Turk does not dispute this.

114. Further, I disagree with Dr. Turk's argument that capturing images of the user's face would render Derakhshani's system inoperable for its intended purpose. (EX2061, ¶¶197, 243.) As I explained above, Derakhshani's system captures images of the user's face for determining the spatial metric, thereby showing it is operable when capturing images of the user's face.

5. Derakhshani and Tanii teach "interpolating the first biometric data and the second biometric data to obtain estimated intermediate biometric data" ('606 Patent, Elements 2[a] and 11[a])

115. Dr. Turk argues that Derakhshani does not teach elements 2[a] and 11[a] because it does not generate a two-dimensional plane. (EX2061, ¶¶204-206.) However, Dr. Turk misunderstands Derakhshani's disclosures. Derakhshani states that "the locations of landmarks (e.g., an eye, nose, mouth, and ear) in the two dimensional image will occur in **a** two dimensional plane." (EX1005, 17:16-18.) "For example, the locations of multiple landmarks may be fit to ***the closest two dimensional plane***." (EX1005, 17:19-22 (emphasis added).) A POSITA would have understood that "a two dimensional plane" refers to any two-dimensional

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

surface such as a two dimensional plane that intersects with the x-axis and y-axis at (5,5), a two-dimensional plane that intersects with the x-axis and the y-axis at (5,4), or a two-dimensional plane that intersects with the x-axis and the y-axis at (4,4). Derakhshani's fitting to the closest two dimensional plane means it determines the points of the two-dimensional plane for generating it. Thus, Dr. Turk is wrong that Derakhshani's interpolation does not generate any planes and that its landmarks are fit to an existing plane.

116. Dr. Turk also argues that there is no motivation to combine Derakhshani and Tanii. (EX2061, ¶206.) As discussed above, Dr. Turk's arguments are incorrect.

6. Derakhshani and Tanii teach "verifying the presence of the user's ears in the at least one first image, and verifying the absence or reduced visibility of the user's ears in the at least one second image" ('606 Patent, Claims 3 and 12)

a. Derakhshani

117. Dr. Turk acknowledges that Derakhshani's captures an image of an ear, but argues that it does not teach verifying the presence, absence, or reduced visibility of the ears. (EX2061, ¶¶207-208.) But Derakhshani states that "a landmark (e.g., an iris, an eye corner, a nose, ***an ear***, or a background object) ***may be identified and located in the plurality of images*** with different focus distances." (EX1005, 16:51-54 (emphasis added).) It also states that "determining 714 the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

spatial metric comprises determining parallax of two or more landmarks that ***appear in a plurality of the images***. . . . This ***parallax effect may be measured*** and used as a spatial metric that reflects the three-dimensional nature of a subject in the view of the sensor." (EX1005, 17:45-55 (emphasis added).) Identifying and locating a landmark or measuring parallax for a landmark that appears in the images means verifying the presence of a landmark, including verifying the presence of an ear in an image. And just like how capturing an image of the user's face is not incidental (see discussion above), capturing an image of the ear is not incidental, as Dr. Turk argues. (EX2061, ¶¶207-208.)

118. Dr. Turk also argues that Derakhshani's use of landmarks for determining the spatial metric undercuts the idea of an ear being absent or having reduced visibility. (EX2061, ¶208.) I disagree. Derakhshani discloses different degrees of focus (a low degree of focus or a high degree of focus) for a landmark. (EX1005, 16:60-17:11.) A landmark having a low degree of focus would be blurry, thereby having reduced visibility. Dr. Turk testified similarly, but used the terms "resolution" and "definition." (EX1059, 142:12-143:21.) Derakhshani explains that "[w]hen the degree is low for a landmark, smaller higher frequency components are expected. By comparing the degree of focus for a landmark in images with different focus distances, the distance from the sensor to the landmark may be estimated." (EX1005, 16:65-17:2.) The distances, including the distance

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

for a blurry landmark, are then used to determine the spatial metric. (EX1005, 17:19-26.) The use of distance information for a landmark shows that Dr. Turk is wrong that Derakhshani merely uses a landmark to estimate the degree of focus. (EX2061, ¶209.) The focus-distance approach involves capturing a first image with a landmark in focus and a second image with the same landmark not in focus (blurry). The spatial metric determination involves comparing the degrees of focus in the first and second images. Dr. Turk confirmed this understanding. (EX1059, 144:11-16, 145:9-12.) Thus, Derakhshani teaches determining the spatial metric of an image of a user's ear having reduced visibility.

b. Tanii

119. Dr. Turk argues that nothing in Tanii suggests verifying the presence or absence of ears in an image. (EX2061, ¶210.) But Tanii's images show otherwise. Figure 3B (below) shows the presence of the user's ear in the first image, and Figure 4B (below) shows the absence of the user's ear in the second image. Further, Dr. Turk's focus on whether Tanii verifies the presence, reduced visibility, or absence of the ears shows his misunderstanding of the Derakhshani-Tanii combination. (EX2061, ¶210.) As discussed above, Derakhshani teaches verifying the user's ears.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Fig.3A

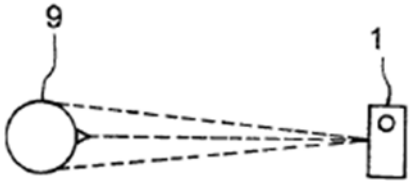


Fig.3B

Fig.4A

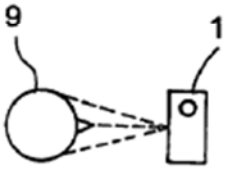
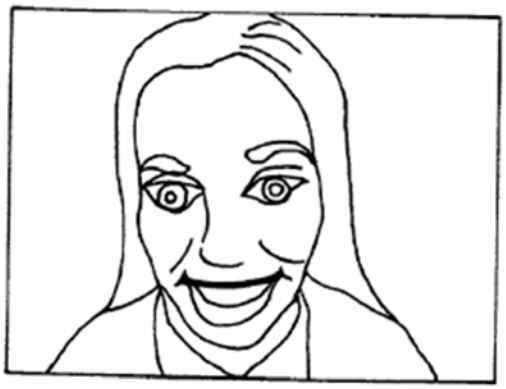


Fig.4B



7. Derakhshani and Tanii teach "determining that the user's face is three-dimensional occurs at a server" ('606 Patent, Claim 14)

120. Dr. Turk's only argument why Derakhshani-Tanii does not teach claim 14 is because Derakhshani captures images of the eyes and not the face. (EX2061, ¶196.) As I discussed above, Dr. Turk is wrong.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

B. Derakhshani, Tanii, And Tahk (Ground 1B)

121. Ground 1B challenges claim 5, which requires "ovals sized on the screen within which the face of the user is placed."

122. Dr. Turk's only argument why Derakhshani-Tanii-Tahk does not teach claim 5 is because Derakhshani captures images of the eyes and not the face. (EX2061, ¶196.) As I discussed above, Dr. Turk is wrong. I note that Dr. Turk does not rebut that Tahk teaches claim 5.

C. Zhang And Tanii (Ground 2A)

1. Zhang and Tanii teach "biometric data" ('606 Patent, Elements 1[b], 10[c], and 19[a])

123. As I explained in ¶¶59-64, Zhang teaches biometric authentication that involves distinguishing live faces from flat surfaces.

124. Dr. Turk argues that Zhang does not teach elements 1[b], 10[c], and 19[a] because its feature points are not "biometric data" used for authentication. (EX2061, ¶¶222-223.) According to Dr. Turk, Zhang's feature points are used only to generate a homography matrix H. (EX2061, ¶223.) This is the same argument that Dr. Turk made for elements 1[f], 10[g], and 19[d] of the '606 patent. I explained why Dr. Turk's arguments are incorrect below. Further, Zhang's facial features meet the description in the '471 specification of "biometric data" (similar to the description in the '606 specification), which Dr. Turk agrees with. (EX1059, 50:11-14.) Zhang states that "feature points are feature points of the face," "such

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

as a corner of an eye, a corner of a mouth, a top of a nose, and so forth." (EX1006, [0027].) Thus, the comparison involves biometric data.

2. Zhang and Tanii teach "comparing[/compare] the first biometric data to[/with] second biometric data to determine whether differences between the at least one first image and the at least one second image match expected differences" ('606 Patent, Elements 1[f], 10[g], and 19[d])

a. Zhang

125. Dr. Turk argues that Zhang does not determine "expected differences." (EX2061, ¶213.) Dr. Turk reasons that color/intensity differences do not relate to expected distortion. (EX2061, ¶¶214-215.) This is false. A POSITA would have understood that color/intensity differences in images indicate depth, thereby captured images having color/intensity differences indicate a three-dimensional object such as a face exhibiting expected distortion. Zhang explains this: "[i]f the image difference meets the threshold value, then the face in the first and second images is determined to be a live face (act 314)." (EX1006, [0034].) Derakhshani explains this concept: "the intensity or color scale can be zoomed in to reveal more subtle changes in the level of the detected image signal. If the captured images are of a ***live eye, it is expected that the range of color or intensity values detected will continue to vary continuously***. In contrast, a spoofed image (e.g., a digital photograph presented to the sensor) may exhibit large discontinuous jumps corresponding to halftones."). (EX1005, 19:30-37.) Additionally, Dr.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Turk's focus on a single pixel in each image is irrelevant because Zhang still determines the differences.

126. I further disagree with Dr. Turk's arguments that the color/intensity differences identified in Zhang does not relate to expected distortion that is based on movement of the camera or the user, and that Zhang makes no prediction of how an image would change due to this movement. (EX2061, ¶214.) This argument mischaracterizes Zhang's teachings and the Petition's Zhang-Tanii combination. Zhang discloses accessing "multiple images including a face presented by a user" and determining whether the face is a "3-dimensional structure or a flat surface" by analyzing differences between images. (EX1006, Abstract, [0025].) Zhang explains that when two images of a flat surface are compared, they can be aligned using a homography matrix such that "the warped image will be similar to the second image" with a "difference between the second image and the warped image [that] is small." (EX1006, [0025].) In contrast, when the images depict a three-dimensional face, "the difference between the second image and the warped image is larger" because the homography matrix cannot align different views of a 3D object. (EX1006, [0025].) The color/intensity differences that Zhang measures are the manifestation of this misalignment, which in turn results from the three-dimensional geometry of the face creating perspective-dependent variations across the images. A POSITA would have

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

understood that these color/intensity differences arise because different parts of the three-dimensional face are at different distances from the camera, causing them to appear differently across viewpoints in ways that a planar transformation cannot correct.

127. Moreover, the Zhang-Tanii combination does not rely on Zhang alone to teach analyzing expected distortion resulting from camera or user movement. Rather, as I explained in my opening declaration, the Zhang-Tanii combination relies on Tanii for teaching that three-dimensional faces exhibit predictable, distance-dependent perspective distortion. (EX1003, ¶¶495-595.) Tanii discloses "where the main object 9 and the cellular phone 1 are close together … an unnatural image results in which the perspective is exaggerated," and explains the technical reason: "the difference between the angle θ1 … and the angle θ2 … decreases as the lens unit 211 comes closer to the main object 9." (EX1007, [0047]-[0049].) Tanii further discloses that "the warp or distortion increases as the angle θ1 … increases" and that "where the main object is close to the imaging portion 2 and the image of the main object occupies a large percentage of the entire image, the exaggeration warp of the image becomes substantial." (Tanii, [0049].) These teachings provide the prediction Dr. Turk claims is missing: as the camera moves closer to a three-dimensional face, the perspective distortion increases in a predictable manner, with "peripheral areas of the main object appear[ing] reduced

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

in size relative to the center area." (EX1007, [0056].) A POSITA would have understood that Zhang's framework would be enhanced by Tanii's teachings that the type and magnitude of differences expected from a three-dimensional face vary predictably with distance due to perspective distortion. The combination teaches comparing observed differences (as Zhang does via color/intensity analysis) to expected differences (as Tanii characterizes via distance-induced distortion) to determine three-dimensionality for verification/evaluation. Dr. Turk's argument that the color/intensity differences do not relate to expected distortion based on movement fails to account for Tanii's contributions to the combination.

128. Dr. Turk alleges that Jumio confuses Zhang's "feature point" and "selected location." (EX2061, ¶217.) As I explained in my opening declaration, the extracted feature points are data that characterize an individual's unique physical characteristics. (EX1003, ¶¶513-514.) The feature point location characterizes the location of the individual's unique physical characteristics. In other words, Zhang's feature point location is part of the feature point data, thereby showing that the mapping is not inconsistent.

129. And Dr. Turk's arguments that the Petition cites to different passages in Zhang is flawed. (EX2061, 110 fn. 3.) Element 1[b] and element 1[f] recite different limitations. Elements 1[b] are directed to processing to obtain first biometric data, while element 1[f] is directed to comparing the first biometric data

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

to the second biometric data. Both sets of passages, [0027] (relied on for element 1[b]) and [0031]-[0032] (relied on for element 1[f]), discuss feature point data.

130. Further, Dr. Turk incorrectly tries to differentiate between feature point data used to generate the homography matrix H and feature point data used in the comparison of images. (EX2061, ¶¶217-221.) However, as Zhang explains, "feature points" are parts of the face, "such as a corner of an eye, a corner of a mouth, a tip of a nose, and so forth" that are extracted from the first and second images. (EX1006, [0027].) The extracted feature points are matched across the first and second images and used to generate the homography matrix H. (EX1006, [0028]-[0029].) Zhang discloses "[t]his matching of the feature points across the first and second images refers to identifying the locations of the same feature points in each of the two images. For example, the locations of the same corner of the same eye in the face in the first and second images are identified." (EX1006, [0028].) Then, when comparing the images, the image difference computes the "difference in color or intensity values at the same pixels (locations) in the two images." (EX1006, [0031].) Zhang uses the same term "locations" when describing the generation of the homography matrix H and also when describing the difference in color/intensity values because it is referring to the same extracted feature points. Using Zhang's example of the corner of the eye as the "location," it would make sense for the comparison between the two images to be a comparison

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

of the same feature point, such as the same corner of the same eye because otherwise a difference in color/intensity value may be due to different features points having different colors or intensity values due to, e.g., the user moving. For example, the color of the corner of an eye is likely to be different than the color of the corner of a mouth, thereby a comparison of the same pixel position having two different feature points (as Dr. Turk suggests) would not be indicative of whether the face is a 3D structure or a flat surface. Thus, Zhang's feature points are used for both generating the homography matrix H and the comparison of the images (part of authenticating the user).

131. Dr. Turk alleges that a three-dimensional image would not align with Zhang's approach and use of a homography matrix, and then concludes that Zhang does not predict how a three-dimensional face would distort across different views. (EX2061, ¶68.) However, Zhang's homography-based approach recognizes the fact that a three-dimensional face produces characteristic differences across views that are not eliminated by planar transformations. Zhang teaches that "two views of a flat (planar) surface are related based on a homography matrix H" such that warping one image using the homography matrix H produces an image similar to the second image. (EX1006, [0025].) If there is a large difference between the first and second images, Zhang determines that "the first and second images include a face that is a live face rather than a flat surface." (EX1006, [0025].) A

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

large difference is due to the images being of a three-dimensional face because different parts of the face are at different depths. The three-dimensional face undergoing different perspective transformations across viewpoints prevents the three-dimensional image from not being well-aligned using the homography matrix, as Turk alleges. (EX1006, [0030]-[0031]; EX2061, ¶68.) And a small difference indicates a flat surface (expected for a 2D photograph). (EX1006, [0025], [0033]-[0034].) A POSITA would have understood that Zhang's approach involves three-dimensional objects (including faces) that exhibit view-dependent distortions that are distinguishable from two-dimensional objects.

132. Moreover, as I explained above and in my opening declaration, the Petition combines Zhang with Tanii's teaching of the distance-induced distortion of a face when the camera comes closer to the face. (EX1003, ¶¶495-505.) Tanii teaches that perspective distortion of three-dimensional faces varies predictably with camera-to-subject distance, with closer distances producing more pronounced "exaggeration warp" where "peripheral areas of the main object appear reduced in size relative to the center area." (EX1007, [0047]-[0049], [0056].) The Zhang-Tanii combination teaches a POSITA that (1) three-dimensional faces produce characteristic differences across images captured at different distances (Zhang's principle) and (2) these differences manifest as distance-induced perspective distortion (Tanii's teaching). Together, the Zhang-Tanii combination teaches

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

determining whether expected distortion exists in the captured images, which is what the ’606 claims require.

b. Tanii

133. Dr. Turk makes arguments about Tanii for the Zhang-Tanii combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii combination. (EX2061, ¶¶224-225, 230.) I explained why Dr. Turk’s arguments are incorrect above.

3. A POSITA would have been motivated to combine Zhang and Tanii

134. Dr. Turk argues that there is no motivation to combine Zhang and Tanii because Tanii destroys Zhang’s “basic objective” that requires distortions to operate. (EX2061, ¶¶243-244, 247.) Dr. Turk’s arguments are however premised on his misunderstanding of the Zhang-Tanii combination. The teachings of Tanii that are part of the Zhang-Tanii combination are its teachings of distance-induced distortions in images of a user’s face. (EX1003, ¶¶495-505) Dr. Turk erroneously bases his arguments on Tanii’s teachings of image correction, which are not part of the Zhang-Tanii combination. (EX2061, ¶¶246, 253-254, 260.)

135. Dr. Turk also criticizes Tanii’s figures for allegedly being inconsistent with its disclosures. (EX2061, ¶257.) Specifically, Dr. Turk alleges that the ears in the warped image (Figure 9) would not “miraculously appear after Tanii’s purported distortion correction.” (EX2061, ¶257.) However, Dr. Turk’s argument

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

shows that he misunderstands Tanii. A POSITA would have understood that Tanii's figures are simplified line drawings that portray abstract representations of the distance-induced distortions caused by the optical effects. (EX1007, [0048], [0049], [0056].) An actual captured image, on the other hand, has more than hundreds of pixels. Thus, Tanii's figures do not depict the level of detail of a captured image. Tanii does not cause ears to "miraculously appear," as Dr. Turk asserts. Instead, the magnification is changing based on the distance from the image center. (EX1003, ¶¶46-51.) Given that Tanii's figures are simplified and abstract depictions, a POSITA would have understood that the ears are present in the actual captured image corresponding to Figure 4B, but because of the magnification, the ears are difficult to see. For example, the rightmost image (below) has ears that are present, but is difficult to see with the human eye due the distance-induced distortion. In those photographs, the user's ears appear increasingly distorted (reduced in lateral size compared with other facial features) as the distance between the sensor and the user is reduced. Also consistent with Tanii's figures, the user's nose exhibits distortion where it is relatively enlarged as this distance is reduced.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

4. Zhang and Tanii teach "interpolating the first biometric data and the second biometric data to obtain estimated intermediate data" ('606 Patent, Elements 2[a] and 11[a])

136. Dr. Turk argues that Zhang does not teach elements 2[a] and 11[a] because it does not model expected distortion. (EX2061, ¶227.) However, the '606 claims do not require modeling expected distortion. Elements 2[a] and 11[a] require interpolating first and second data to obtain estimated intermediate data, which Zhang teaches. As I explained in my opening declaration, Zhang teaches additional intermediate images existing between the two images being compared for authentication. (EX1003, ¶¶535-539; Zhang, [0035].) The intermediate images are used for feature point extraction and feature point matching. (EX1006, [0036].) A POSITA would have been motivated to generate predictions of what the intermediate images should look like, thereby interpolating the data of the two images to obtain estimated intermediate data.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

5. Zhang and Tanii teach "verifying the presence of the user's ears in the at least one first image, and verifying the absence or reduced visibility of the user's ears in the at least one second image" ('606 Patent, Claims 3 and 12)

137. Dr. Turk makes arguments about Tanii for the Zhang-Tanii combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii combination. (EX2061, ¶¶228-229.) As I explained above, Tanii teaches claims 3 and 12.

A. Tanii Is Analogous Art

1. Tanii is from the same field of endeavor

138. Dr. Turk argues that Tanii is not in the same field of endeavor as the '606 patent because the '606 uses distortion and Tanii tries to remove it as part of its image correction process. (EX2061, ¶¶231-232.) But Dr. Turk's argument that Tanii assumes distortion exists and corrects for it is misplaced. The Derakhshani-Tanii combination and the Zhang-Tanii combination rely on Tanii for its teachings of distance-induced distortion, not on its teachings of an image correction process.

139. Further, a POSITA would not have understood Tanii's teachings to be so narrow as Dr. Turk argues. Instead, a POSITA would have understood that one of Tanii's teachings is related to identifying and accounting for distance-induced distortions of three-dimensional faces in images captured using a camera.

140. Thus, both the '606 patent claims and Tanii relate to distance-induced distortion. That is, Tanii is directed to the same function (accounting of distance-

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

induced distortions) and structure (camera), and as a result is in the same field of endeavor as the ’606 patent.

2. Tanii is reasonably pertinent to the problems faced by the inventors of the ’606 Patent

141. Dr. Turk acknowledges that one of the problems the ’606 inventors were faced with is “this long-understood issue” of a three-dimensional face exhibiting distortions when the camera and face are close together during image capture. (EX2061, ¶237.) Dr. Turk does not dispute that Tanii is reasonably pertinent to this distance-induced distortion problem, but instead argues that Tanii’s teachings would not be helpful because the problem was known. (EX2061, ¶237.) But whether Tanii’s teachings were known does not determine whether Tanii is reasonably pertinent.

142. I note that the Petition stated that “[t]he ’606 Patent describes . . . methods to distinguish real. three-dimensional faces from two-dimensional pictures of a face during a facial-authentication process.” (Petition, 3.) The Petition did not admit that distinguishing a three-dimensional face from a two-dimensional picture of a face was a problem faced by the ’606 patent, as Dr. Turk alleges. (EX2061, ¶236.)

143. Dr. Turk argues that the ’606 patent solves the spoofing problem and Tanii does not. (EX2061, ¶¶238-239.) But spoofing is not the only alleged

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

problem faced by the ’606 inventors. As discussed above, Tanii solves the ’606 problem related to distance-induced distortions.

144. Dr. Turk criticizes Tanii because allegedly “the ’606 patent uses the distortion that Tanii tries to remove.” (EX2061, ¶232.) He argues that “Tanii works against the challenged claims, by removing the very thing that enabled the patent’s inventor to solve the spoofing problem.” (EX2061, ¶241.) I disagree. A POSITA would have understood that Tanii’s teachings of distance-induced distortion is suitable for use for user authentication.

145. And Dr. Turk’s argument about purpose is flawed. (EX2061, ¶242.) As described in the context of the relevant art is discussed, [0006] states “[i]f the image of the face of a person or the like is captured from a point that is relatively nearby, an image is obtained in which the perspective appears exaggerated. . . . When a camera comes close to a three-dimensional object that has depth, the peripheral areas of the object, which are farther away from the camera than the front area, appear to be closer to the front area than they really are, resulting in a warp that causes the image to appear as if the perspective were exaggerated (the type of warp in which the perspective becomes exaggerated, as in this case, will hereinafter be termed ‘exaggeration warp’, as distinguished from distortion).” (EX1007, [0006].) Thus, Tanii accounts for warp that causes the image to appear as exaggerated.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

V. ’938 PATENT

A. Derakhshani And Tanii (Ground 1A)

1. Derakhshani and Tanii teach “a human face” (’938 Patent, Element 1[pre]), “user’s face” (’938 Patent, Elements 8[pre]/20[a]), and “a head of the person” (’938 Patent, Element 17[a])

a. Derakhshani

146. As I explained in ¶¶9-16, Derakhshani teaches biometric authentication that involves determining whether images depict a three-dimensional subject (user).

147. Dr. Turk argues that Derakhshani does not teach capturing an image of the user’s face. (EX2061, ¶¶262, 263.) I disagree.

148. Dr. Turk does not dispute that Derakhshani teaches capturing at least part of a user’s face. He however argues that capturing part of the user’s face (other than the eye) is incidental. (EX2061, ¶277.) I do not agree that capturing part of the user’s face is incidental.

149. Derakhshani discloses determining a spatial metric “based on a distance from a sensor to a landmark that appears in a plurality of the images each having a different respective focus distance.” (EX1005, 16:44-47.) The determined spatial metric is used to determine three-dimensionality: “[a] high value for this spatial metric may indicate a three-dimensional subject and a higher likelihood that the subject is a live eye, while a low value may indicate a higher

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

likelihood that the subject is a two-dimensional spoof attack." (EX1005, 17:22-26.)

150. Derakhshani specifically names "a nose, an ear," and a "mouth" as example landmarks identified in the captured images. (EX1005, 16:51-54, 17:16-17.) The passage at 17:29-36 discusses "landmarks on a particular user's face." (EX1005, 17:29-36.) And the passage at 3:13-15 states that the "landmark may be a portion of a face depicted in the images." (EX1005, 3:13-15.) Derakhshani also states that an image may comprise multiple "landmarks (e.g., an eye, nose, mouth, and ear." (EX1005, 17:16-19; *see also* 3:5-7 ("two or more landmarks that appear in a plurality of the images"), 17:2-6 ("multiple landmarks").) Dr. Turk also agreed that Derakhshani captures an image having multiple landmarks. (EX1059, 136:25-137:3.) It would have been obvious for an image comprising an eye, nose, mouth, and ear, as Derakhshani discloses, to be an image of a face because these are facial features that when they are all included in an image, it would be an image of a user's face. (EX1005, 17:16-20.)

151. Derakhshani also names a "background object" as an example landmark. (EX1005, 16:51-54.) It would have been obvious for an image comprising an eye and a background object (landmark) to be an image of a user's face. For example, the background object may be a painting on a wall located behind the user. In order to capture the user's eye and the painting in the image,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

the sensor would have to be back far enough from the user such that the user's face would be captured. Thus, Dr. Turk's suggestion (at, *e.g.*, ¶¶270 and 274 of his declaration (EX2061)) that Derakhshani captures images of the user's eyes only is wrong. I also note that Derakhshani does not disclose capturing the user's eye vasculature as its primary purpose, as Dr. Turk argues at ¶¶272 and 277 of his declaration.

152. Derakhshani further states that its authentication is "based ***in part*** on one or more images of the white of an eye 410." (EX1005, 5:65-67.) Being based "in part" means that some of the captured images are of the white of an eye 410 (such as the images captured using the sensor 6-12 cm from the eye (EX1005, 13:35-40)), and some of the captured images are not just of the white of an eye, such as images of the user's face.

153. Dr. Turk's declaration argument that the spatial metric determination relies on characteristics of the user's eye only is also incorrect. (EX2061, ¶271 (citing EX1005, 4:19-21).) Derakhshani's spatial metric involves capturing a plurality of images at difference distances. (EX1005, 16:44-47, 17:45-47.) One image of the plurality of captured images may be a close-up image, and another image of the plurality of captured images may be captured at a different distance further away from the user. The close-up image may capture the eye vasculature

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

(with enough precision), and the further-away image may capture the landmark(s) (with enough precision).

154. The passage at 4:19-21 that Dr. Turk cites describes image(s) used for determining the match score. The match score is determined using image(s) that capture the vasculature of the eye and are "subjected to further analysis" (step 616) ***after and separate from*** the determination of the spatial metric (step 714, which is part of step 710 performed as step 604). (EX1005, 14:59-64, 16:44-18:7.) As Dr. Turk testified, the eye vasculature is determined after determining the liveness score and determining the liveness score is separate from determining the match score. (EX1059, 130:3-7, 132:9-11.) Further, Dr. Turk agreed during his deposition that the spatial metric determination is not solely based on characteristics of the user's eye. (EX1059, 126:4-13.) Dr. Turk also agreed that Derakhshani captures images having more than the user's eye. (EX1059, 136:16-24.)

155. Further, a POSITA would have understood a prior art reference based on not only what is depicted in the figures, but also the teachings discussed in the specification. Even Dr. Turk looks to the specification when reviewing a prior art reference. (EX1059, 16:3-17:24.) Accordingly, Dr. Turk's declaration arguments that Derakhshani captures images of the eyes only because that is what is depicted in the figures is without merit because Derakhshani's specification suggests

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

otherwise. (EX2061, ¶¶262, 272-273; EX1005, 16:44-54, 17:16-17.) Dr. Turk also testified that if there is an embodiment with the image including a nose, ear, mouth, and eye, it would be an image that is not depicted in Derakhshani figures. (EX1059, 137:4-17.) Indeed, Derakhshani discloses such an embodiment, and thus, its figures do not show all of the disclosed embodiments. (EX1005, 17:16-19.)

156. Dr. Turk argues that Derakhshani's system is not adapted to capture the user's face or head. (EX2061, ¶¶272, 274.) I disagree. To make an accurate determination of the spatial metric, a POSITA would have understood that there must be a discernable difference in the degree of focus (for the focus-distance approach) or movement of the landmarks (for the parallax approach), which would mean there must be a discernable difference in distance between the user and the sensor. This would mean at least one captured image would be taken such that the sensor is far enough away from the user, thereby including more than just the user's eye in the image. Derakhshani discusses this at 17:55-59: "[i]f all the landmarks in the image undergo the ***same apparent displacement*** due to relative motion of the sensor, ***i.e., the difference in the parallax effect for the landmarks is small***, then the subject viewed by the camera has [a] higher likelihood of being a two-dimensional spoof attack." (EX1005, 17:55-59.) A POSITA would have understood from this passage that a large difference in the parallax effect for the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

landmarks would be needed when the object is three-dimensional in order to avoid a false determination of a two-dimensional spoof attack.

157. Further, Derakhshani teaches achieving a finer resolution by using a reduced dynamic range for the sensor. "[O]ne or more images may be captured using a reduced dynamic range for the sensor (e.g., a camera) so that a finer resolution in intensity of detected light is achieved in a range within which it occurs in the captured images. In this manner, the intensity or color scale can be zoomed in to reveal more subtle changes in the level of the detected image signal." (EX1005, 19:24-32.) Thereby, by adjusting the parameters, the same sensor may be used to capture images of the vasculature of an eye with a finer resolution and images of the face. This means Derakhshani's system is adapted to capture the user's head/face.

158. Dr. Turk also argues that simultaneous capture of the eye vasculature and face in the same image with enough precision for both is required. (EX2061, ¶276.) I disagree based on Derakhshani's disclosures of capturing a plurality of images for the spatial metric determination. For the focus-distance approach, the plurality of images have different respective focus distances for a given landmark. Derakhshani provides an example where the landmarks in an image include "an eye, nose, mouth, and ear." (EX1005, 17:16-17.) It explains that a landmark may be "in focus" or may be blurred (also described as "the degree of focus is low").

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

(EX1005, 16:56-66.) A POSITA would have understood Derakhshani's disclosures of different focus distances in different images and a landmark being in focus and blurry means the landmark is in focus in one image and blurred in another image. The spatial metric is determined based on distances of the landmarks. (EX1005, 16:66-17:6, 17:12-22, 17:49-52.) A high value for the spatial metric indicates a three-dimensional subject and a low value indicates a two-dimensional spoof attack. (EX1005, 17:22-26.) A POSITA would have understood that large estimated distances from the landmarks would occur when some of the landmarks are blurred (out of focus) and result in a high value for the spatial metric. Thus, Derakhshani does not need to capture images that simultaneously capture both the vasculature of an eye and a face with enough precision in order to determine three-dimensionality or authentication, as Dr. Turk argues. (EX2061, ¶276.)

159. Instead, a POSITA would have understood that Derakhshani may be able to determine three-dimensionality when not simultaneously capturing both the vasculature of an eye and face with enough precision. For example, Derakhshani may capture a close-up image with the user's eye in focus but the ear is not in focus, and may also capture another image that is further away from the user where the ear is in focus. The spatial metric may be determined using the close-up image and the further-away image, both having a discernable difference in distance of the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

ear landmark, thereby resulting in a high spatial metric and determination of three-dimensionality. The match score may be determined using the close-up image. In such a situation, an image that simultaneously captures both the eye vasculature and the face with enough precision is not required. Thus, not every image needs to be close enough to capture the vasculature of the eye with sufficient clarity, as Dr. Turk argues. (EX2061, ¶¶263, 267.)

b. Tanii

160. Tanii also teaches capturing images of a user's three-dimensional face involving distance-induced distortions, as shown in Figures 3A-3B and 4A-4B (reproduced below). (EX1007, Abstract ("an image is captured with the user's face at a close range"), [0006], [0013], [0047]-[0049], [0061]-[0063].) Dr. Turk does not dispute this, but rather argues that Tanii does not identify a face. (EX2061, ¶283.) This is incorrect. Tanii states that it "detects the size of the image of the main object." (EX1007, [0074].) Tanii explains that "the image of the main object, particularly the image of the face of the user who is holding the cellular phone 1, is captured." (EX1007, [0061].) Thus, Tanii detects the size of the user's face. In order to detect the size of the user's face, it must first detect the user's face. Further, Tanii describes different ways to detect the user's face, such as "extracting the contours of the main object using image processing." (EX1007, [0069].) Tanii also describes "identif[ying] the region of the image of the main

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

object based on the location of clear edges in the image as well as on the color distribution in the image." (EX1007, [0076].)

Fig.3A

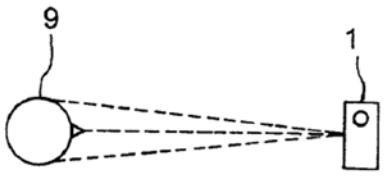


Fig.3B

Fig.4A

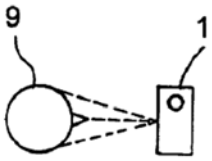
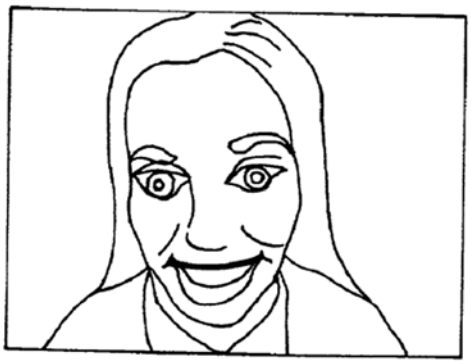


Fig.4B



2. Derakhshani and Tanii teach "biometric data" ('938 Patent, Elements 1[a] and 17[d])

161. Dr. Turk argues that Derakhshani does not teach elements 1[a] and 17[d] of the '938 patent because its landmarks are not "biometric data." First, I disagree with Dr. Turk's separation of landmarks into what he calls "image

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

landmarks" and "anatomical landmarks." (EX2061, ¶58.) Dr. Turk seems to create this separation because "image landmarks" are used for the liveness score determination and "anatomical landmarks" are used for the match score determination. (EX2061, ¶58.) However, Derakhshani teaches the eye as being an example landmark used for both the liveness score determination and the match score determination, thereby showing that a landmark cannot be classified as either an "image landmark" or an "anatomical landmark." (EX1005, 14:25-29; 17:16-19.)

162. Dr. Turk argues that Derakhshani's landmark is not "biometric data" because it may be an iris, eye corner, or ear, which is not used to identify a user. (EX2061, ¶282.) But Derakhshani uses the iris, eye corner, or ear to determine a liveness score or a match score, both of which are used to authenticate the user. (EX1005, 11:46-51, 16:44-54, 17:12-26, 19:16-18, 21:8-28, Figs. 6-8.)

163. I note that Dr. Turk's requirement provided in his declaration that "biometric data" be used to identify a user is inconsistent with the '938 specification, which states: "[t]he term biometric data may include among other information biometric information concerning facial features and ***path parameters***. ***Examples of path parameters may include an acceleration and speed of the mobile device, angle of the mobile device during image capture, distance of the mobile device to the user, path direction in relation to the user's face position in***

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

relation to the user, or any other type parameter associated with movement of the mobile device or the user face in relation to a camera. Other data may also be included such as GPS data, device identification information, and the like." ('471 patent (EX1001), 6:40-50.) This is inconsistent with his testimony where he agreed with the specification's description of "biometric data." (EX1059, 50:11-14.)

3. Derakhshani and Tanii teach "the second biometric data does not have expected differences compared to the first biometric data, … the expected differences result from fish-eye type distortion" ('938 Patent, Element 1[d]), "when the first data and the second data do not have expected differences indicating the user exhibits three-dimensionality" ('938 Patent, Element 8[g]), "determine whether expected differences are not present" ('938 Patent, Element 17[d]), and "determining the user does not exhibit the expected degree of three-dimensionality" ('938 Patent, Element 20[h])

164. Tanii teaches the "expected differences," "fish-eye type distortion," and "expected degree of three-dimensionality" of elements 1[d], 8[g], 17[d], and 20[h]. (EX1007, Figs. 3A-3B, 4A-4B, 5, [0047]-[0049].) Tanii explains that when there is a sufficient distance between the main object 9 and the cellular phone 1, the image appears natural. (EX1007, [0047], Figs. 3A-3B.) When the main object 9 and cellular phone 1 are too close to each other, the perspective is exaggerated resulting in an unnatural image. (EX1007, [0047], Figs. 4A-4B.) This exaggeration in perspective and unnaturalness is an expected distortion because it

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

occurs due to known and expected differences that result from the light ray striking different points of the main object 9. (EX1007, [0048]-[0049], Fig. 5.)

165. Dr. Turk does not dispute these teachings. Instead, he focuses his arguments on whether Tanii evaluates degree of distance-induced distortion, or measures, assesses, or undoes three-dimensional distortion. (EX2061, ¶¶265, 285.) But the '938 claims do not require measuring or undoing three-dimensional distortion. Further, the Derakhshani-Tanii combination relies on Derakhshani for its teachings of authenticating the user, which Dr. Turk does not dispute Derakhshani teaches.

166. Moreover, Dr. Turk is incorrect that Tanii does not assess three-dimensional distortion and that it assumes an image requires correction. (EX2061, ¶335.) Tanii determines the distance from the main object to the camera. (EX1007, [0101].) It discloses "where the distance to the main object equals or exceeds the threshold value D1, it is determined that the main object and the imaging portion 2 are located a sufficient distance apart and that as a result no correction is needed." (EX1007, [0101].) That is, when there is no distance-induced distortion, Tanii does not perform a correction. Tanii also teaches determining the degree of distance-induced distortion and applying the appropriate amount of correction. It discloses "[w]here the distance to the main object is less than the threshold value D1" or "less than the threshold value D2," then "a low

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

degree of correction" or "a high degree of correction" is needed. (EX1007, [0101].) It further discloses "[i]t can be automatically determined whether or not correction is needed." (EX1007, [0136], Claim 7.) In other words, Tanii identifies and assesses distance-induced distortions.

167. Dr. Turk also argues that Derakhshani compares the captured images to an expected plane/shape, but does not compare data of two captured images together. (EX2061, ¶269.) This is incorrect. Dr. Turk points to 3:1-6 and 17:27-36 to support his argument, but these passages describe one embodiment. There are other passages in Derakhshani, such as 16:51-54 and 17:45-47, that teach comparing data of two captured images together. The 16:51-54 passage and surrounding text describe the spatial metric being determined based on "a distance from a sensor to a landmark that ***appears in a plurality of the images*** each having a different respective focus distance." (EX1005, 16:44-54 (emphasis added).) It states that "the focus distance may be adjusted for different images." (*Id.*; *see also* 16:44-54 ("the plurality of the images with different focus distances").) The 17:45-47 passage and surrounding text describe "parallax of two or more landmarks that ***appear in a plurality of the images***," where the "plurality of images [are] taken from different perspectives on the subject" and "the sensor is moved about the subject during image capture." (EX1005, 17:45-64 (emphasis

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

added).) Thus, Derakhshani captures a plurality of images having different distances, and the data from the images are compared to each other.

168. Dr. Turk further argues that Tanii aesthetically corrects images and states that its alleged "process of enlargement" does not make up for the deficiencies of Derakhshani and is incompatible with Derakhshani. (EX2061, ¶¶283, 285.) However, Dr. Turk misunderstands the Derakhshani-Tanii combination which does not use Tanii's aesthetic/warp correction. The Derakhshani-Tanii combination includes Tanii's teachings of distance-induced distortions in images of a user's face. (EX1003, ¶¶603-611.) Further, whether Tanii operates the same way when the image does not include a face is irrelevant because Tanii teaches capturing and detecting the user's face. (EX2061, ¶284; EX1007, [0062].)

4. A POSITA would have been motivated to combine Derakhshani and Tanii

169. Dr. Turk argues that there is no motivation to combine Derakhshani and Tanii because Tanii destroys Derakhshani's "basic objective" that requires distortions to operate. (EX2061, ¶¶332-333, 335.) Dr. Turk's arguments are however premised on his misunderstanding of the Derakhshani-Tanii combination and the Petition's motivation to combine. The teachings of Tanii that are part of the Derakhshani-Tanii combination are its teachings of distance-induced

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

distortions in images of a user's face—not teachings of image correction. (EX1003, ¶¶603-611; EX2061, ¶165.)

170. Additionally, Dr. Turk's argument that images at different distances having different distortion is not a motivation to combine again shows that he misunderstands the Derakhshani-Tanii combination. (EX2061, ¶¶286, 286.) As I explained in my opening declaration, Tanii teaches an obvious alternative to Derakhshani's focus-distance and parallax approaches. (EX1003, ¶¶603, 607.) Derakhshani's approaches require a camera with extreme sensitivity, not suitable for mobile phones that have wide angle lenses to capture a wide field of view. (EX1003, ¶608.) I note that Dr. Turk acknowledges that Derakhshani's approach has limited applicability. (EX2061, ¶337.) Dr. Turk also acknowledges that Tanii's teachings apply in situations where images are captured using a wide-angle lens. (EX2061, ¶338.) Thus, a POSITA would have been motivated to use Tanii's method when using a mobile device that has a wide-angle lens to capture the user's image(s). (EX1003, ¶¶609-611.) Dr. Turk fails to rebut this motivation.

171. I disagree with Dr. Turk's argument that Derakhshani and Tanii are incompatible. (EX2061, ¶338.) Dr. Turk's argues that this alleged incompatibility is due to Derakhshani and Tanii having different imaging requirements. (EX2061, ¶338.) Dr. Turk argues that Derakhshani needs to capture the eye vasculature with sufficient clarity and the camera must be positioned very close to the user's eye,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

eliminating the ability to capture the user's face or the ability to use Tanii's correction method designed for capturing images at an arms length. This argument misunderstands Derakhshani's disclosures and the Derakhshani-Tanii combination. Derakhshani's system captures a plurality of images at different conditions. As I explained above, in Derakhshani, different images may be used by different functions of its authentication process. For the spatial metric determination (part of the liveness detection), Derakhshani uses images captured at different focus distances, or from different perspectives to analyze landmarks such as "a nose, an ear," "a mouth," or a "background object." (EX1005, 16:51-54, 17:16-19.) A POSITA would have understood that capturing an image containing facial landmarks results in capturing an image of the user's face. For the match score determination (eye vasculature comparison), Derakhshani uses close-up images to resolve the eye vasculature with sufficient precision. (EX1005, 13:35-40.) A POSITA would have understood that Derakhshani's multi-image framework accommodates both types of captured mages: some images optimized for the spatial metric determination (which may include facial features at varying distances from the sensor), and some images optimized for vasculature matching (which may include images of the eye vasculature). Dr. Turk assumes that all images must simultaneously capture both the eye vasculature and the facial landmarks with precision, but Derakhshani does not disclose or suggest such a

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

requirement. Further, Dr. Turk's argument about incompatibility supports the motivation to combine because Derakhshani's approach is limited and applying Tanii's teachings would result in improving the implementation of biometric authentication in a mobile device as Derakhshani already envisions. (EX1005, 5:23-26; EX1003, ¶¶607-611.) Derakhshani's focus-distance and parallax approaches use specialized camera capabilities that may not be suitable for mobile phones with wide-angle lenses. Tanii teaches capturing images of the user's face using a cellular phone with a wide-angle lens at an arms length. (EX1007, [0007], [0047].) Tanii discloses that when "the user's face is close to the camera" with "a wide-angle focal length," "the exaggeration warp becomes more noticeable." (EX1007, [0007].) As I explained my opening declaration, a POSITA would have wanted to implement Derakhshani's authentication process on a mobile device, and would have been motivated to use Tanii's teachings about distance-induced distortion of an image of a user's face as an alternative or supplemental way to determining spatial metric. Derakhshani teaches using "a reduced dynamic range for the sensor . . . so that a finer resolution in intensity of detected light is achieved," demonstrating that the same sensor can be adjusted to capture different types of images by modifying camera parameters. (EX1005, 19:24-32.) A POSITA would have understood that similar adjustments would allow Derakhshani's system to capture both close-up images for the match score

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

determination and further away images for the spatial metric determination (using Derakhshani's and/or Tanii's approach). As I explained in my opening declaration, a POSITA would have known how to apply Tanii's teachings to Derakhshani's system using well-known optics principles relating to camera design, and there would have been a reasonable expectation of success. (EX1003, ¶610.) Dr. Turk does not dispute this.

172. Further, I disagree with Dr. Turk's argument that capturing images of the user's face would render Derakhshani's system inoperable for its intended purpose. (EX2061, ¶¶280, 332.) As I explained above, Derakhshani's system captures images of the user's face for determining the spatial metric, thereby showing it is operable when capturing images of the user's face.

5. Derakhshani and Tanii teach "compar[ing] at least portions of the first data, second data, or both to enrollment data" ('938 Patent, Element 5[a])

173. Derakhshani teaches identifying landmarks and using the "relative positions of landmarks on a particular user's face . . . [to] generate a three-dimensional model that is stored as part of a reference record." (EX1005, 17:32-36.) The features in the image (e.g., of the eye) are extracted and compared to the reference record. (EX1005, 4:19-24, 7:19-34, 9:59-67, 13:62-14:9.) In other words, Derakhshani uses both terms "landmarks" and "features" to describe the data of the reference record, thereby Derakhshani does not make a distinction

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

between "landmarks" and "features," as Dr. Turk argues. (EX2061, ¶¶292-296.) And "landmarks" are also used for the spatial metric determination, thereby the data for liveness detection is the same as the data used for the reference record (what Dr. Turk refers to as "enrollment records" (EX2061, ¶296)).

174. Dr. Turk's argument that "landmarks" and "features" not synonymous because the "landmarks" used for liveness determination are separate from the "features" used for authentication is incorrect. (EX2061, ¶¶295-296.) As discussed above, Derakhshani determines the liveness score as part of the authentication process, and the spatial metric is determined as part of determining the liveness score. The passage at 8:60-67 explains that the liveness score is determined as part of Derakhshani's authentication process: "authentication module 540 may receive image data from the requesting user's computing device … and may process the image data to authenticate the user. In some implementations, the ***authentication module may determine liveness scores*** for images received from a user and may accept or reject the images based on the liveness scores." (EX1005, 8:60-67 (emphasis added).) Also, the passage at 14:42-45 states "[i]f the ***liveness score*** indicates a high likelihood that live eye is depicted in the images, the one or more images are accepted 616 and subjected to further analysis to ***complete the authentication process.***" "complet[ing] the authentication process" means the liveness score is part of the authentication

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

process. Further, Derakhshani names the eye as a landmark, and the eye is also a feature. (EX1005, 17:16-17 ("***landmarks (e.g., an eye,*** nose, mouth, and ear).)

6. Derakhshani and Tanii teach "the first biometric data and the second biometric data comprise [at least in part] image data of facial features" ('938 Patent, Claims 7 and 24)

175. Dr. Turk argues that Derakhshani's landmarks are not image data of facial features since it does not capture images of the user's face. (EX2061, ¶299.) As I discussed above, Derakhshani does capture images of faces, the images may include facial feature landmarks such as "an eye, nose, mouth, and ear." (EX1005, 17:16-19.)

B. Zhang And Tanii (Ground 2A)

1. Zhang and Tanii teach "biometric data" ('938 Patent, Elements 1[a] and 17[d])

176. As I explained in ¶¶59-64, Zhang teaches biometric authentication that involves distinguishing live faces from flat surfaces.

177. Dr. Turk argues that Zhang does not teach elements 1[a] and 17[d] because its feature points are not "biometric data" used for authentication. (EX2061, ¶¶315-316.) According to Dr. Turk, Zhang's feature points are used only to generate a homography matrix H. (EX2061, ¶316.) This is the same argument that Dr. Turk made for elements 1[d], 8[g], 17[d], and 20[h] of the '938 patent. I explained why Dr. Turk's arguments are incorrect below. Further, Zhang's facial features meet the description in the '471 specification of "biometric

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

data" (similar to the description in the '938 specification), which Dr. Turk agrees with. (EX1059, 50:11-14.) Zhang states that "feature points are feature points of the face," "such as a corner of an eye, a corner of a mouth, a top of a nose, and so forth." (EX1006, [0027].) Thus, the comparison involves biometric data.

2. Zhang and Tanii teach "the second biometric data does not have expected differences compared to the first biometric data, … the expected differences result from fish-eye type distortion" ('938 Patent, Element 1[d]), "when the first data and the second data do not have expected differences indicating the user exhibits three-dimensionality" ('938 Patent, Element 8[g]), "determine whether expected differences are not present" ('938 Patent, Element 17[d]), and "determining the user does not exhibit the expected degree of three-dimensionality" ('938 Patent, Element 20[h])

a. Zhang

178. Dr. Turk argues that Zhang does not determine "expected differences" or "expected distortion." (EX2061, ¶¶303-304.) Dr. Turk reasons that color/intensity differences do not relate to expected distortion. (EX2061, ¶¶304-306.) This is false. A POSITA would have understood that color/intensity differences in images indicate depth, thereby captured images having color/intensity differences indicate a three-dimensional object such as a face exhibiting expected distortion. Zhang explains this: "[i]f the image difference meets the threshold value, then the face in the first and second images is determined to be a live face (act 314)." (EX1006, [0034].) Derakhshani explains

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

this concept: "the intensity or color scale can be zoomed in to reveal more subtle changes in the level of the detected image signal. If the captured images are of a ***live eye, it is expected that the range of color or intensity values detected will continue to vary continuously***. In contrast, a spoofed image (e.g., a digital photograph presented to the sensor) may exhibit large discontinuous jumps corresponding to halftones."). (EX1005, 19:30-37.) Additionally, Dr. Turk's focus on a single pixel in each image is irrelevant because Zhang still determines the differences.

179. I further disagree with Dr. Turk's arguments that the color/intensity differences identified in Zhang does not relate to expected distortion that is based on movement of the camera or the user, and that Zhang makes no prediction of how an image would change due to this movement. (EX2061, ¶304.) This argument mischaracterizes Zhang's teachings and the Petition's Zhang-Tanii combination. Zhang discloses accessing "multiple images including a face presented by a user" and determining whether the face is a "3-dimensional structure or a flat surface" by analyzing differences between images. (EX1006, Abstract, [0025].) Zhang explains that when two images of a flat surface are compared, they can be aligned using a homography matrix such that "the warped image will be similar to the second image" with a "difference between the second image and the warped image [that] is small." (EX1006, [0025].) In contrast, when

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

the images depict a three-dimensional face, "the difference between the second image and the warped image is larger" because the homography matrix cannot align different views of a 3D object. (EX1006, [0025].) The color/intensity differences that Zhang measures are the manifestation of this misalignment, which in turn results from the three-dimensional geometry of the face creating perspective-dependent variations across the images. A POSITA would have understood that these color/intensity differences arise because different parts of the three-dimensional face are at different distances from the camera, causing them to appear differently across viewpoints in ways that a planar transformation cannot correct.

180. Moreover, the Zhang-Tanii combination does not rely on Zhang alone to teach analyzing expected distortion resulting from camera or user movement. Rather, as I explained in my opening declaration, the Zhang-Tanii combination relies on Tanii for teaching that three-dimensional faces exhibit predictable, distance-dependent perspective distortion. (EX1003, ¶¶740-750.) Tanii discloses "where the main object 9 and the cellular phone 1 are close together … an unnatural image results in which the perspective is exaggerated," and explains the technical reason: "the difference between the angle θ1 … and the angle θ2 … decreases as the lens unit 211 comes closer to the main object 9." (EX1007, [0047]-[0049].) Tanii further discloses that "the warp or distortion increases as the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

angle θ1 … increases" and that "where the main object is close to the imaging portion 2 and the image of the main object occupies a large percentage of the entire image, the exaggeration warp of the image becomes substantial." (Tanii, [0049].) These teachings provide the prediction Dr. Turk claims is missing: as the camera moves closer to a three-dimensional face, the perspective distortion increases in a predictable manner, with "peripheral areas of the main object appear[ing] reduced in size relative to the center area." (EX1007, [0056].) A POSITA would have understood that Zhang's framework would be enhanced by Tanii's teachings that the type and magnitude of differences expected from a three-dimensional face vary predictably with distance due to perspective distortion. The combination teaches comparing observed differences (as Zhang does via color/intensity analysis) to expected differences (as Tanii characterizes via distance-induced distortion) to determine three-dimensionality for authentication. Dr. Turk's argument that the color/intensity differences do not relate to expected distortion based on movement fails to account for Tanii's contributions to the combination.

181. Dr. Turk alleges that Jumio confuses Zhang's "feature point" and "selected location." (EX2061, ¶307.) As I explained in my opening declaration, the extracted feature points are data that characterize an individual's unique physical characteristics. (EX1003, ¶¶757-758.) The feature point location characterizes the location of the individual's unique physical characteristics. In

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

other words, Zhang's feature point location is part of the feature point data, thereby showing that the mapping is not inconsistent.

182. And Dr. Turk's arguments that the Petition cites to different passages in Zhang is flawed. (EX2061, 151 fn. 4.) Elements 1[a] and 1[c] of the '938 Patent recite different limitations. Element 1[a] is directed to receiving or deriving first biometric data, while element 1[c] is directed to comparing the first biometric data with the second biometric data for expected differences. Both sets of passages, [0027] (relied on for element 1[a]) and [0031]-[0032] (relied on for element 1[c]), discuss feature point data.

183. Further, Dr. Turk incorrectly tries to differentiate between feature point data used to generate the homography matrix H and feature point data used in the comparison of images. (EX2061, ¶¶307-311.) However, as Zhang explains, "feature points" are parts of the face, "such as a corner of an eye, a corner of a mouth, a tip of a nose, and so forth" that are extracted from the first and second images. (EX1006, [0027].) The extracted feature points are matched across the first and second images and used to generate the homography matrix H. (EX1006, [0028]-[0029].) Zhang discloses "[t]his matching of the feature points across the first and second images refers to identifying the locations of the same feature points in each of the two images. For example, the locations of the same corner of the same eye in the face in the first and second images are identified." (EX1006,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

[0028].) Then, when comparing the images, the image difference computes the "difference in color or intensity values at the same pixels (locations) in the two images." (EX1006, [0031].) Zhang uses the same term "locations" when describing the generation of the homography matrix H and also when describing the difference in color/intensity values because it is referring to the same extracted feature points. Using Zhang's example of the corner of the eye as the "location," it would make sense for the comparison between the two images to be a comparison of the same feature point, such as the same corner of the same eye because otherwise a difference in color/intensity value may be due to different features points having different colors or intensity values due to, e.g., the user moving. For example, the color of the corner of an eye is likely to be different than the color of the corner of a mouth, thereby a comparison of the same pixel position having two different feature points (as Dr. Turk suggests) would not be indicative of whether the face is a 3D structure or a flat surface. Thus, Zhang's feature points are used for both generating the homography matrix H and the comparison of the images (part of authenticating the user).

184. Dr. Turk alleges that a three-dimensional image would not align with Zhang's approach and use of a homography matrix, and then concludes that Zhang does not predict how a three-dimensional face would distort across different views. (EX2061, ¶68.) However, Zhang's homography-based approach recognizes the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

fact that a three-dimensional face produces characteristic differences across views that are not eliminated by planar transformations. Zhang teaches that "two views of a flat (planar) surface are related based on a homography matrix H" such that warping one image using the homography matrix H produces an image similar to the second image. (EX1006, [0025].) If there is a large difference between the first and second images, Zhang determines that "the first and second images include a face that is a live face rather than a flat surface." (EX1006, [0025].) A large difference is due to the images being of a three-dimensional face because different parts of the face are at different depths. The three-dimensional face undergoing different perspective transformations across viewpoints prevents the three-dimensional image from not being well-aligned using the homography matrix, as Turk alleges. (EX1006, [0030]-[0031]; EX2061, ¶68.) And a small difference indicates a flat surface (expected for a 2D photograph). (EX1006, [0025], [0033]-[0034].) A POSITA would have understood that Zhang's approach involves three-dimensional objects (including faces) that exhibit view-dependent distortions that are distinguishable from two-dimensional objects.

185. Moreover, as I explained above and in my opening declaration, the Petition combines Zhang with Tanii's teaching of the distance-induced distortion of a face when the camera comes closer to the face. (EX1003, ¶¶740-750.) Tanii teaches that perspective distortion of three-dimensional faces varies predictably

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

with camera-to-subject distance, with closer distances producing more pronounced "exaggeration warp" where "peripheral areas of the main object appear reduced in size relative to the center area." (EX1007, [0047]-[0049], [0056].) The Zhang-Tanii combination teaches a POSITA that (1) three-dimensional faces produce characteristic differences across images captured at different distances (Zhang's principle) and (2) these differences manifest as distance-induced perspective distortion (Tanii's teaching). Together, the Zhang-Tanii combination teaches determining whether expected distortion exists in the captured images, which is what the '938 claims require.

b. Tanii

186. Dr. Turk makes arguments about Tanii for the Zhang-Tanii combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii combination. (EX2061, ¶¶317-319.) I explained why Dr. Turk's arguments are incorrect above.

3. A POSITA would have been motivated to combine Zhang and Tanii

187. Dr. Turk argues that there is no motivation to combine Zhang and Tanii because Tanii destroys Zhang's "basic objective" that requires distortions to operate. (EX2061, ¶¶332-333, 336.) Dr. Turk's arguments are however premised on his misunderstanding of the Zhang-Tanii combination. The teachings of Tanii that are part of the Zhang-Tanii combination are its teachings of distance-induced

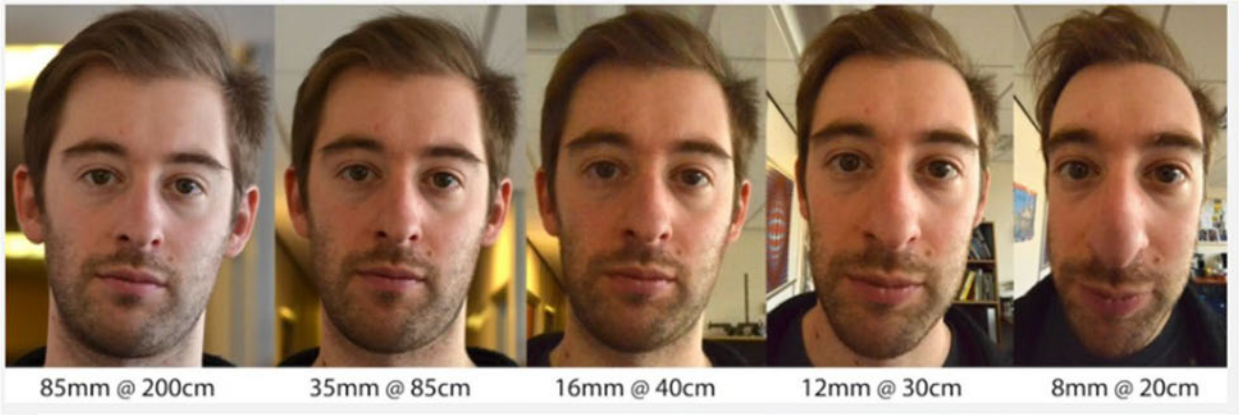
Jumio EX1060
Jumio v. FaceTec IPR2025-00109

distortions in images of a user's face. (EX1003, ¶¶740-750.) Dr. Turk erroneously bases his arguments on Tanii's teachings of image correction, which are not part of the Zhang-Tanii combination. (EX2061, ¶¶335, 342-343, 349-350.)

188. Dr. Turk also criticizes Tanii's figures for allegedly being inconsistent with its disclosures. (EX2061, ¶346.) Specifically, Dr. Turk alleges that the ears in the warped image (Figure 9) would not "miraculously appear after Tanii's purported distortion correction." (EX2061, ¶346.) However, Dr. Turk's argument shows that he misunderstands Tanii. A POSITA would have understood that Tanii's figures are simplified line drawings that portray abstract representations of the distance-induced distortions caused by the optical effects. (EX1007, [0048], [0049], [0056].) An actual captured image, on the other hand, has more than hundreds of pixels. Thus, Tanii's figures do not depict the level of detail of a captured image. Tanii does not cause ears to "miraculously appear," as Dr. Turk asserts. Instead, the magnification is changing based on the distance from the image center. (EX1003, ¶¶46-51.) Given that Tanii's figures are simplified and abstract depictions, a POSITA would have understood that the ears are present in the actual captured image corresponding to Figure 4B, but because of the magnification, the ears are difficult to see. For example, the rightmost image (below) has ears that are present, but is difficult to see with the human eye due the distance-induced distortion. In those photographs, the user's ears appear

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

increasingly distorted (reduced in lateral size compared with other facial features) as the distance between the sensor and the user is reduced. Also consistent with Tanii's figures, the user's nose exhibits distortion where it is relatively enlarged as this distance is reduced.



C. Tanii Is Analogous Art

1. Tanii is from the same field of endeavor

189. Dr. Turk argues that Tanii is not in the same field of endeavor as the '938 patent because the '938 uses distortion and Tanii tries to remove it as part of its image correction process. (EX2061, ¶¶.320-321.) But Dr. Turk's argument that Tanii assumes distortion exists and corrects for it is misplaced. The Derakhshani-Tanii combination and the Zhang-Tanii combination rely on Tanii for its teachings of distance-induced distortion, not on its teachings of an image correction process.

190. Further, a POSITA would not have understood Tanii's teachings to be so narrow as Dr. Turk argues. Instead, a POSITA would have understood that one

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

of Tanii's teachings is related to identifying and accounting for distance-induced distortions of three-dimensional faces in images captured using a camera.

191. Thus, both the '938 patent claims and Tanii relate to distance-induced distortion. That is, Tanii is directed to the same function (accounting of distance-induced distortions) and structure (camera), and as a result is in the same field of endeavor as the '938 patent.

2. Tanii is reasonably pertinent to the problems faced by the inventors of the '938 Patent

192. Dr. Turk acknowledges that one of the problems the '938 inventors were faced with is "this long-understood issue" of a three-dimensional face exhibiting distortions when the camera and face are close together during image capture. (EX2061, ¶326.) Dr. Turk does not dispute that Tanii is reasonably pertinent to this distance-induced distortion problem, but instead argues that Tanii's teachings would not be helpful because the problem was known. (EX2061, ¶326.) But whether Tanii's teachings were known does not determine whether Tanii is reasonably pertinent.

193. I note that the Petition stated that "[t]he '938 Patent describes . . . methods to distinguish real, three-dimensional faces from two-dimensional pictures of a face during a facial-authentication process." (Petition, 4.) The Petition did not admit that distinguishing a three-dimensional face from a two-dimensional picture

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

of a face was a problem faced by the ’938 patent, as Dr. Turk alleges. (EX2061, ¶325.)

194. Dr. Turk argues that the ’938 patent solves the spoofing problem and Tanii does not. (EX2061, ¶¶327-328.) But spoofing is not the only alleged problem faced by the ’938 inventors. As discussed above, Tanii solves the ’938 problem related to distance-induced distortions.

195. Dr. Turk criticizes Tanii because allegedly “the ’938 patent uses the distortion that Tanii tries to remove.” (EX2061, ¶321.) He argues that “Tanii works against the challenged claims, by removing the very thing that enabled the patent’s inventor to solve the spoofing problem.” (EX2061, ¶330.) I disagree. A POSITA would have understood that Tanii’s teachings of distance-induced distortion is suitable for use for user authentication.

196. And Dr. Turk’s argument about purpose is flawed. (EX2061, ¶331.) As described in the context of the relevant art is discussed, [0006] states “[i]f the image of the face of a person or the like is captured from a point that is relatively nearby, an image is obtained in which the perspective appears exaggerated. . . . When a camera comes close to a three-dimensional object that has depth, the peripheral areas of the object, which are farther away from the camera than the front area, appear to be closer to the front area than they really are, resulting in a warp that causes the image to appear as if the perspective were exaggerated (the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

type of warp in which the perspective becomes exaggerated, as in this case, will hereinafter be termed 'exaggeration warp', as distinguished from distortion)." (EX1007, [0006].) Thus, Tanii accounts for warp that causes the image to appear as exaggerated.

VI. '910 PATENT

A. Derakhshani And Tanii (Ground 1A)

1. Derakhshani and Tanii teach "verifying[/evaluating] three-dimensionality of a[/the] user" ('910 Patent, Elements 1[pre], 10[pre], and 22[pre])

a. Derakhshani

197. As I explained in ¶¶9-16, Derakhshani teaches biometric authentication that involves determining whether images depict a three-dimensional subject (user).

198. Dr. Turk agreed that Derakhshani teaches authenticating three-dimensionality of a user. (EX1059, 128:21-129:1.) But he argues that Derakhshani does not teach capturing an image of the user's face. (EX2061, ¶¶351, 357.) However, none of the independent claims of the '910 patent requires capturing an image of the user's face, or authenticating the user using an image of the user's face. Rather, the independent claims of the '910 patent require verifying or evaluating three-dimensionality of a user, which, as I explained in my opening declaration, the Derakhshani-Tanii combination teaches. (EX1003, ¶¶886-887 (citing EX1005, 1:11-25, 2:4-30, 3:14-15, 5:22-27, 6:3-5, 9:10-22, 16:44-18:4).)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

199. Notably, Dr. Turk does not dispute that Derakhshani teaches capturing at least part of a user's face. He however argues that capturing part of the user's face (other than the eye) is incidental. (EX2061, ¶366.) I do not agree that capturing part of the user's face is incidental.

200. Derakhshani discloses determining a spatial metric "based on a distance from a sensor to a landmark that appears in a plurality of the images each having a different respective focus distance." (EX1005, 16:44-47.) The determined spatial metric is used to determine three-dimensionality: "[a] high value for this spatial metric may indicate a three-dimensional subject and a higher likelihood that the subject is a live eye, while a low value may indicate a higher likelihood that the subject is a two-dimensional spoof attack." (EX1005, 17:22-26.)

201. Derakhshani specifically names "a nose, an ear," and a "mouth" as example landmarks identified in the captured images. (EX1005, 16:51-54, 17:16-17.) The passage at 17:29-36 discusses "landmarks on a particular user's face." (EX1005, 17:29-36.) And the passage at 3:13-15 states that the "landmark may be a portion of a face depicted in the images." (EX1005, 3:13-15.) Derakhshani also states that an image may comprise multiple "landmarks (e.g., an eye, nose, mouth, and ear." (EX1005, 17:16-19; *see also* 3:5-7 ("two or more landmarks that appear in a plurality of the images"), 17:2-6 ("multiple landmarks").) Dr. Turk also

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

agreed that Derakhshani captures an image having multiple landmarks. (EX1059, 136:25-137:3.) It would have been obvious for an image comprising an eye, nose, mouth, and ear, as Derakhshani discloses, to be an image of a face because these are facial features that when they are all included in an image, it would be an image of a user's face. (EX1005, 17:16-20.)

202. Derakhshani also names a "background object" as an example landmark. (EX1005, 16:51-54.) It would have been obvious for an image comprising an eye and a background object (landmark) to be an image of a user's face. For example, the background object may be a painting on a wall located behind the user. In order to capture the user's eye and the painting in the image, the sensor would have to be back far enough from the user such that the user's face would be captured. Thus, Dr. Turk's suggestion (at, *e.g.*, ¶¶359 and 363 of his declaration (EX2061)) that Derakhshani captures images of the user's eyes only is wrong. I also note that Derakhshani does not disclose capturing the user's eye vasculature as its primary purpose, as Dr. Turk argues at ¶¶361 and 366 of his declaration.

203. Derakhshani further states that its authentication is "based ***in part*** on one or more images of the white of an eye 410." (EX1005, 5:65-67 (emphasis added).) Being based "in part" means that some of the captured images are of the white of an eye 410 (such as the images captured using the sensor 6-12 cm from

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

the eye (EX1005, 13:35-40)), and some of the captured images are not just of the white of an eye, such as images of the user's face.

204. Dr. Turk's declaration argument that the spatial metric determination relies on characteristics of the user's eye only is also incorrect. (EX2061, ¶360 (citing EX1005, 4:19-21).) Derakhshani's spatial metric involves capturing a plurality of images at difference distances. (EX1005, 16:44-47, 17:45-47.) One image of the plurality of captured images may be a close-up image, and another image of the plurality of captured images may be captured at a different distance further away from the user. The close-up image may capture the eye vasculature (with enough precision), and the further-away image may capture the landmark(s) (with enough precision).

205. The passage at 4:19-21 that Dr. Turk cites describes image(s) used for determining the match score. The match score is determined using image(s) that capture the vasculature of the eye and are "subjected to further analysis" (step 616) ***after and separate from*** the determination of the spatial metric (step 714, which is part of step 710 performed as step 604). (EX1005, 11:42-45, 14:59-64, 16:44-18:7.) As Dr. Turk testified, the eye vasculature is determined after determining the liveness score and determining the liveness score is separate from determining the match score. (EX1059, 130:3-7, 132:9-11.) Further, Dr. Turk agreed during his deposition that the spatial metric determination is not solely based on

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

characteristics of the user's eye. (EX1059, 126:4-13.) Dr. Turk also agreed that Derakhshani captures images having more than the user's eye. (EX1059, 136:16-24.)

206. Further, a POSITA would have understood a prior art reference based on not only what is depicted in the figures, but also the teachings discussed in the specification. Even Dr. Turk looks to the specification when reviewing a prior art reference. (EX1059, 16:3-17:24.) Accordingly, Dr. Turk's declaration arguments that Derakhshani captures images of the eyes only because that is what is depicted in the figures is without merit because Derakhshani's specification suggests otherwise. (EX2061, ¶¶351, 361-362; EX1005, 16:44-54, 17:16-17.) Dr. Turk also testified that if there is an embodiment with the image including a nose, ear, mouth, and eye, it would be an image that is not depicted in Derakhshani figures. (EX1059, 137:4-17.) Indeed, Derakhshani discloses such an embodiment, and thus, its figures do not show all of the disclosed embodiments. (EX1005, 17:16-19.)

207. Dr. Turk argues that Derakhshani's system is not adapted to capture the user's face. (EX2061, ¶¶361, 363.) I disagree. To make an accurate determination of the spatial metric, a POSITA would have understood that there must be a discernable difference in the degree of focus (for the focus-distance approach) or movement of the landmarks (for the parallax approach), which would

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

mean there must be a discernable difference in distance between the user and the sensor. This would mean at least one captured image would be taken such that the sensor is far enough away from the user, thereby including more than just the user's eye in the image. Derakhshani discusses this at 17:55-59: "[i]f all the landmarks in the image undergo the ***same apparent displacement*** due to relative motion of the sensor, ***i.e., the difference in the parallax effect for the landmarks is small***, then the subject viewed by the camera has [a] higher likelihood of being a two-dimensional spoof attack." (EX1005, 17:55-59 (emphasis added).) A POSITA would have understood from this passage that a large difference in the parallax effect for the landmarks would be needed when the object is three-dimensional in order to avoid a false determination of a two-dimensional spoof attack.

208. Further, Derakhshani teaches achieving a finer resolution by using a reduced dynamic range for the sensor. "[O]ne or more images may be captured using a reduced dynamic range for the sensor (e.g., a camera) so that a finer resolution in intensity of detected light is achieved in a range within which it occurs in the captured images. In this manner, the intensity or color scale can be zoomed in to reveal more subtle changes in the level of the detected image signal." (EX1005, 19:24-32.) Thereby, by adjusting the parameters, the same sensor may be used to capture images of the vasculature of an eye with a finer resolution and

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

images of the face. This means Derakhshani's system is adapted to capture the user's head/face.

209. Dr. Turk also argues that simultaneous capture of the eye vasculature and face in the same image with enough precision for both is required. (EX2061, ¶365.) I disagree based on Derakhshani's disclosures of capturing a plurality of images for the spatial metric determination. For the focus-distance approach, the plurality of images have different respective focus distances for a given landmark. Derakhshani provides an example where the landmarks in an image include "an eye, nose, mouth, and ear." (EX1005, 17:16-17.) It explains that a landmark may be "in focus" or may be blurred (also described as "the degree of focus is low"). (EX1005, 16:57-66.) A POSITA would have understood Derakhshani's disclosures of different focus distances in different images and a landmark being in focus and blurry means the landmark is in focus in one image and blurred in another image. The spatial metric is determined based on distances of the landmarks. (EX1005, 16:66-17:6, 17:12-22, 17:49-52.) A high value for the spatial metric indicates a three-dimensional subject and a low value indicates a two-dimensional spoof attack. (EX1005, 17:22-26.) A POSITA would have understood that large estimated distances from the landmarks would occur when some of the landmarks are blurred (out of focus) and result in a high value for the spatial metric. Thus, Derakhshani does not need to capture images that

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

simultaneously capture both the vasculature of an eye and a face with enough precision in order to determine three-dimensionality or authentication, as Dr. Turk argues. (EX2061, ¶365.)

210. Instead, a POSITA would have understood that Derakhshani may be able to determine three-dimensionality when not simultaneously capturing both the vasculature of an eye and face with enough precision. For example, Derakhshani may capture a close-up image with the user's eye in focus but the ear is not in focus, and may also capture another image that is further away from the user where the ear is in focus. The spatial metric may be determined using the close-up image and the further-away image, both having a discernable difference in distance of the ear landmark, thereby resulting in a high spatial metric and determination of three-dimensionality. The match score may be determined using the close-up image. In such a situation, an image that simultaneously captures both the eye vasculature and the face with enough precision is not required. Thus, not every image needs to be close enough to capture the vasculature of the eye with sufficient clarity, as Dr. Turk argues. (EX2061, ¶¶352, 356.)

b. Tanii

211. Tanii also teaches capturing images of a user's three-dimensional face involving distance-induced distortions, as shown in Figures 3A-3B and 4A-4B (reproduced below). (EX1007, Abstract ("an image is captured with the user's

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

face at close range"), [0006], [0013], [0047]-[0049], [0061]-[0063].) Dr. Turk does not dispute this, but rather argues that Tanii does not identify a face. (EX2061, ¶371.) This is incorrect. Tanii states that it "detects the size of the image of the main object." (EX1007, [0074].) Tanii explains that "the image of the main object, particularly the image of the face of the user who is holding the cellular phone 1, is captured." (EX1007, [0061].) Thus, Tanii detects the size of the user's face. In order to detect the size of the user's face, it must first detect the user's face. Further, Tanii describes different ways to detect the user's face, such as "extracting the contours of the main object using image processing." (EX1007, [0069].) Tanii also describes "identif[ying] the region of the image of the main object based on the location of clear edges in the image as well as on the color distribution in the image." (EX1007, [0076].)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Fig.3A

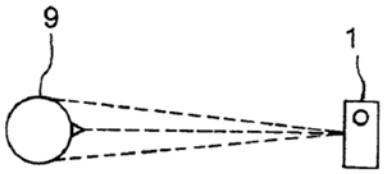


Fig.3B

Fig.4A

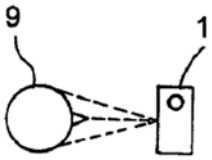
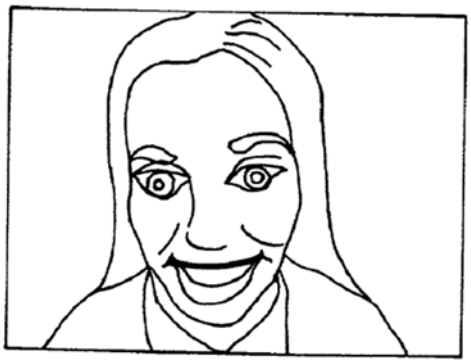


Fig.4B



2. Derakhshani and Tanii teach "comparing the first data to the second data to determine whether expected differences exist" ('910 Patent, Elements 1[d5] and 10[f]) and "comparing one or more aspects of the user's head … to determine whether expected differences, between the first image and the second image, exist" ('910 Patent, Element 22[d])

212. Tanii teaches the "expected differences" of elements 1[d5], 10[f], and 22[d]. (EX1007, Figs. 3A-3B, 4A-4B, 5, [0047]-[0049].) Tanii explains that when there is a sufficient distance between the main object 9 and the cellular phone 1,

the image appears natural. (EX1007, [0047], Figs. 3A-3B.) When the main object 9 and cellular phone 1 are too close to each other, the perspective is exaggerated resulting in an unnatural image. (EX1007, [0047], Figs. 4A-4B.) This exaggeration in perspective and unnaturalness is an expected distortion because it occurs due to known and expected differences that result from the light ray striking different points of the main object 9. (EX1007, [0048]-[0049], Fig. 5.)

213. Dr. Turk does not dispute these teachings. Instead, he focuses his arguments on whether Tanii evaluates degree of distance-induced distortion, or measures, assesses, or undoes three-dimensional distortion. (EX2061, ¶¶354, 373.) But the '910 claims do not require evaluating the degree of distance-induced distortion, or measuring or undoing three-dimensional distortion. Further, the Derakhshani-Tanii combination relies on Derakhshani for its teachings of authenticating the user, which Dr. Turk does not dispute Derakhshani teaches.

214. Moreover, Dr. Turk is incorrect that Tanii does not assess three-dimensional distortion and that it assumes an image requires correction. (EX2061, ¶439.) Tanii determines the distance from the main object to the camera. (EX1007, [0101].) It discloses "where the distance to the main object equals or exceeds the threshold value D1, it is determined that the main object and the imaging portion 2 are located a sufficient distance apart and that as a result no correction is needed." (EX1007, [0101].) That is, when there is no distance-

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

induced distortion, Tanii does not perform a correction. Tanii also teaches determining the degree of distance-induced distortion and applying the appropriate amount of correction. It discloses "[w]here the distance to the main object is less than the threshold value D1" or "less than the threshold value D2," then "a low degree of correction" or "a high degree of correction" is needed. (EX1007, [0101].) It further discloses "[i]t can be automatically determined whether or not correction is needed." (EX1007, [0136], Claim 7.) In other words, Tanii identifies and assesses distance-induced distortions.

215. Dr. Turk also argues that Derakhshani compares the captured images to an expected plane/shape, but does not compare data of two captured images together. (EX2061, ¶358.) This is incorrect. Dr. Turk points to 3:1-6 and 17:27-36 to support his argument, but these passages describe one embodiment. There are other passages in Derakhshani, such as 16:51-54 and 17:45-47, that teach comparing data of two captured images together. The 16:51-54 passage and surrounding text describe the spatial metric being determined based on "a distance from a sensor to a landmark that ***appears in a plurality of the images*** each having a different respective focus distance." (EX1005, 16:44-54 (emphasis added).) It states that "the focus distance may be adjusted for different images." (*Id.*; *see also* 16:44-54 ("the plurality of the images with different focus distances").) The 17:45-47 passage and surrounding text describe "parallax of two or more

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

landmarks that ***appear in a plurality of the images***," where the "plurality of images [are] taken from different perspectives on the subject" and "the sensor is moved about the subject during image capture." (EX1005, 17:45-64 (emphasis added).) Thus, Derakhshani captures a plurality of images having different distances, and the data from the images are compared to each other.

216. Dr. Turk further argues that Tanii aesthetically corrects images and states that its alleged "process of enlargement" does not make up for the deficiencies of Derakhshani and is incompatible with Derakhshani. (EX2061, ¶¶371, 373.) However, Dr. Turk misunderstands the Derakhshani-Tanii combination which does not use Tanii's aesthetic/warp correction. The Derakhshani-Tanii combination includes Tanii's teachings of distance-induced distortions in images of a user's face. (EX1003, ¶¶877-885.) Further, whether Tanii operates the same way when the image does not include a face is irrelevant because Tanii teaches capturing and detecting the user's face. (EX2061, ¶372; EX1007, [0062].)

3. A POSITA would have been motivated to combine Derakhshani and Tanii

217. Dr. Turk argues that there is no motivation to combine Derakhshani and Tanii because Tanii destroys Derakhshani's "basic objective" that requires distortions to operate. (EX2061, ¶¶436-437, 440.) Dr. Turk's arguments are however premised on his misunderstanding of the Derakhshani-Tanii combination

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

and the Petition's motivation to combine. The teachings of Tanii that are part of the Derakhshani-Tanii combination are its teachings of distance-induced distortions in images of a user's face—not teachings of image correction. (EX1003, ¶¶877-885; EX2061, ¶439.)

218. Additionally, Dr. Turk's argument that images at different distances having different distortion is not a motivation to combine again shows that he misunderstands the Derakhshani-Tanii combination. (EX2061, ¶¶374.) As I explained in my opening declaration, Tanii teaches an obvious alternative to Derakhshani's focus-distance and parallax approaches. (EX1003, ¶¶877, 881.) Derakhshani's approaches require a camera with extreme sensitivity, not suitable for mobile phones that have wide angle lenses to capture a wide field of view. (EX1003, ¶882.) I note that Dr. Turk acknowledges that Derakhshani's approach has limited applicability. (EX2061, ¶441.) Dr. Turk also acknowledges that Tanii's teachings apply in situations where images are captured using a wide-angle lens. (EX2061, ¶442.) Thus, a POSITA would have been motivated to use Tanii's method when using a mobile device that has a wide-angle lens to capture the user's image(s). (EX1003, ¶¶883-885.) Dr. Turk fails to rebut this motivation.

219. I disagree with Dr. Turk's argument that Derakhshani and Tanii are incompatible. (EX2061, ¶442.) Dr. Turk's argues that this alleged incompatibility is due to Derakhshani and Tanii having different imaging requirements. (EX2061,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

¶442.) Dr. Turk argues that Derakhshani needs to capture the eye vasculature with sufficient clarity and the camera must be positioned very close to the user's eye, eliminating the ability to capture the user's face or the ability to use Tanii's correction method designed for capturing images at an arms length. This argument misunderstands Derakhshani's disclosures and the Derakhshani-Tanii combination. Derakhshani's system captures a plurality of images at different conditions. As I explained above, in Derakhshani, different images may be used by different functions of its authentication process. For the spatial metric determination (part of the liveness detection), Derakhshani uses images captured at different focus distances, or from different perspectives to analyze landmarks such as "a nose, an ear," "a mouth," or a "background object." (EX1005, 16:51-54, 17:16-19.) A POSITA would have understood that capturing an image containing facial landmarks results in capturing an image of the user's face. For the match score determination (eye vasculature comparison), Derakhshani uses close-up images to resolve the eye vasculature with sufficient precision. (EX1005, 13:35-40.) A POSITA would have understood that Derakhshani's multi-image framework accommodates both types of captured mages: some images optimized for the spatial metric determination (which may include facial features at varying distances from the sensor), and some images optimized for vasculature matching (which may include images of the eye vasculature). Dr. Turk assumes that all

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

images must simultaneously capture both the eye vasculature and the facial landmarks with precision, but Derakhshani does not disclose or suggest such a requirement. Further, Dr. Turk's argument about incompatibility supports the motivation to combine because Derakhshani's approach is limited and applying Tanii's teachings would result in improving the implementation of biometric authentication in a mobile device as Derakhshani already envisions. (EX1005, 5:23-26; EX1003, ¶¶877-885.) Derakhshani's focus-distance and parallax approaches use specialized camera capabilities that may not be suitable for mobile phones with wide-angle lenses. Tanii teaches capturing images of the user's face using a cellular phone with a wide-angle lens at an arms length. (EX1007, [0007], [0047].) Tanii discloses that when "the user's face is close to the camera" with "a wide-angle focal length," "the exaggeration warp becomes more noticeable." (EX1007, [0007].) As I explained my opening declaration, a POSITA would have wanted to implement Derakhshani's authentication process on a mobile device, and would have been motivated to use Tanii's teachings about distance-induced distortion of an image of a user's face as an alternative or supplemental way to determining spatial metric. Derakhshani teaches using "a reduced dynamic range for the sensor . . . so that a finer resolution in intensity of detected light is achieved," demonstrating that the same sensor can be adjusted to capture different types of images by modifying camera parameters. (EX1005, 19:24-32.) A

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

POSITA would have understood that similar adjustments would allow Derakhshani's system to capture both close-up images for the match score determination and further away images for the spatial metric determination (using Derakhshani's and/or Tanii's approach). As I explained in my opening declaration, a POSITA would have known how to apply Tanii's teachings to Derakhshani's system using well-known optics principles relating to camera design, and there would have been a reasonable expectation of success. (EX1003, ¶884.) Dr. Turk does not dispute this.

220. Further, I disagree with Dr. Turk's argument that capturing images of the user's face would render Derakhshani's system inoperable for its intended purpose. (EX2061, ¶¶369, 436.) As I explained above, Derakhshani's system captures images of the user's face for determining the spatial metric, thereby showing it is operable when capturing images of the user's face.

4. Derakhshani and Tanii teach "verifying the presence of one or more features[/ears] … in the at least one first image, and verifying the absence or reduced visibility of the one or more features [/ears] … in the at least one second image" ('910 Patent, Claims 3 and 12)

a. Derakhshani

221. Dr. Turk acknowledges that Derakhshani's captures an image of an ear, but argues that it does not teach verifying the presence, absence, or reduced visibility of the ears. (EX2061, ¶¶378-379.) But Derakhshani states that "a

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

landmark (e.g., an iris, an eye corner, a nose, ***an ear***, or a background object) ***may be identified and located in the plurality of images*** with different focus distances." (EX1005, 16:51-54 (emphasis added).) It also states that "determining 714 the spatial metric comprises determining parallax of two or more landmarks that ***appear in a plurality of the images***. . . . This ***parallax effect may be measured*** and used as a spatial metric that reflects the three-dimensional nature of a subject in the view of the sensor." (EX1005, 17:45-55 (emphasis added).) Identifying and locating a landmark or measuring parallax for a landmark that appears in the images means verifying the presence of a landmark, including verifying the presence of an ear in an image. And just like how capturing an image of the user's face is not incidental (see discussion above), capturing an image of the ear is not incidental, as Dr. Turk argues. (EX2061, ¶¶378-379.)

222. Dr. Turk also argues that Derakhshani's use of landmarks for determining the spatial metric undercuts the idea of an ear being absent or having reduced visibility. (EX2061, ¶379.) I disagree. Derakhshani discloses different degrees of focus (a low degree of focus or a high degree of focus) for a landmark. (EX1005, 16:60-17:11.) A landmark having a low degree of focus would be blurry, thereby having reduced visibility. Dr. Turk testified similarly, but used the terms "resolution" and "definition." (EX1059, 142:12-143:21.) Derakhshani explains that "[w]hen the degree is low for a landmark, smaller higher frequency

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

components are expected. By comparing the degree of focus for a landmark in images with different focus distances, the distance from the sensor to the landmark may be estimated." (EX1005, 16:65-17:2.) The distances, including the distance for a blurry landmark, are then used to determine the spatial metric. (EX1005, 17:19-26.) The use of distance information for a landmark shows that Dr. Turk is wrong that Derakhshani merely uses a landmark to estimate the degree of focus. (EX2061, ¶380.) The focus-distance approach involves capturing a first image with a landmark in focus and a second image with the same landmark not in focus (blurry). The spatial metric determination involves comparing the degrees of focus in the first and second images. Dr. Turk confirmed this understanding. (EX1059, 144:11-16, 145:9-12.) Thus, Derakhshani teaches determining the spatial metric of an image of a user's ear having reduced visibility.

b. Tanii

223. Dr. Turk argues that nothing in Tanii suggests verifying the presence or absence of ears in an image. (EX2061, ¶381.) But Tanii's images show otherwise. Figure 3B (below) shows the presence of the user's ear in the first image, and Figure 4B (below) shows the absence of the user's ear in the second image. Further, Dr. Turk's focus on whether Tanii verifies the presence, reduced visibility, or absence of the ears shows his misunderstanding of the Derakhshani-

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Tanii combination. (EX2061, ¶381.) As discussed above, Derakhshani teaches verifying the user's ears.

Fig.3A

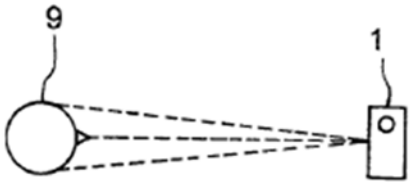


Fig.3B

Fig.4A

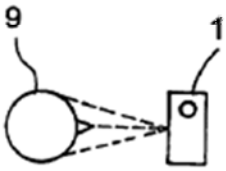
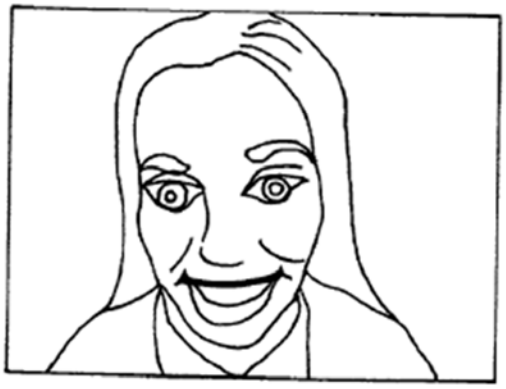


Fig.4B



5. Derakhshani and Tanii teach "comparing the first data, second data, or both to enrollment data" ('910 Patent, Element 5[a])

224. Derakhshani teaches identifying landmarks and using the "relative positions of landmarks on a particular user's face . . . [to] generate a three-

dimensional model that is stored as part of a reference record." (EX1005, 17:32-36.) The features in the image (e.g., of the eye) are extracted and compared to the reference record. (EX1005, 4:19-24, 7:19-34, 9:59-67, 13:62-14:9.) In other words, Derakhshani uses both terms "landmarks" and "features" to describe the data of the reference record, thereby Derakhshani does not make a distinction between "landmarks" and "features," as Dr. Turk argues. (EX2061, ¶¶387-391.) And "landmarks" are also used for the spatial metric determination, thereby the data for liveness detection is the same as the data used for the reference record (what Dr. Turk refers to as "enrollment records" (EX2061, ¶391)).

225. Dr. Turk's argument that "landmarks" and "features" are not synonymous because the "landmarks" used for liveness determination are separate from the "features" used for authentication is incorrect. (EX2061, ¶¶390-391.) As discussed above, Derakhshani determines the liveness score as part of the authentication process, and the spatial metric is determined as part of determining the liveness score. The passage at 8:60-67 explains that the liveness score is determined as part of Derakhshani's authentication process: "authentication module 540 may receive image data from the requesting user's computing device … and may process the image data to authenticate the user. In some implementations, the ***authentication module may determine liveness scores*** for images received from a user and may accept or reject the images based on the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

liveness scores." (EX1005, 8:60-67 (emphasis added).) Also, the passage at 11:42-45 states "[i]f the ***liveness score*** indicates a high likelihood that live eye is depicted in the images, the one or more images are accepted 616 and subjected to further analysis to ***complete the authentication process.***" (EX1005, 11:42-45 (emphasis added).) "[C]omplet[ing] the authentication process" means the liveness score is part of the authentication process. Further, Derakhshani names the eye as a landmark, and the eye is also a feature. (EX1005, 17:16-17 ("***landmarks (e.g., an eye,*** nose, mouth, and ear) (emphasis added).)

6. Derakhshani and Tanii teach "biometric data" ('910 Patent, Claims 7 and 16) and "the first data and the second data comprise a map[ping] of facial features" ('910 Patent, Claims 8 and 17)

226. Dr. Turk argues that Derakhshani does not teach claims 7, 8, 16, and 17 of the '910 patent because its landmarks are not "biometric data." Dr. Turk's first reason is because Derakhshani teaches "image landmarks," which may be a background object. (EX2061, ¶395.) But this is incorrect. First, I disagree with Dr. Turk's separation of landmarks into what he calls "image landmarks" and "anatomical landmarks." (EX2061, ¶58.) Dr. Turk seems to create this separation because "image landmarks" are used for the liveness score determination and "anatomical landmarks" are used for the match score determination. (EX2061, ¶58.) However, Derakhshani teaches the eye as being an example landmark used for both the liveness score determination and the match score determination,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

thereby showing that a landmark cannot be classified as either an "image landmark" or an "anatomical landmark." (EX1005, 14:25-29; 17:16-19.) Further, even if a background object does not contain "biometric data," this does not mean all "image landmarks" cannot contain biometric data. Dr. Turk even testified that his meaning of "biometric data" is "data about a human, human biology or body or something." (EX1059, 44:4-13.) Dr. Turk cites 15:51-54 of Derakhshani, which provides "an iris, an eye corner, a nose, an ear, or a background object" as example landmarks. (EX2061, ¶387 (citing EX1005, 16:51-54).) Because an image of an iris, an eye corner, a nose, and an ear may have data about human biology or body, then Derakhshani's so-called landmarks may contain "biometric data."

227. Dr. Turk's second reason is because Derakhshani's landmark may be an iris, eye corner, or ear, which is not used to identify a user. (EX2061, ¶395.) But Derakhshani uses the iris, eye corner, or ear to determine a liveness score or a match score, both of which are used to authenticate the user. (EX1005, 11:46-51, 16:44-54, 17:12-26, 19:16-18, 21:8-28, Figs. 6-8.)

228. I note that Dr. Turk's requirement provided in his declaration that "biometric data" be used to identify a user is inconsistent with the '910 specification, which states: "[t]he term biometric data may include among other information biometric information concerning facial features (distorted or undistorted) and ***path parameters***. ***Examples of path parameters may include an***

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

acceleration and speed of the mobile device, angle of the mobile device during image capture, distance of the mobile device to the user, path direction in relation to the user's face position in relation to the user, or any other type parameter associated with movement of the mobile device or the user face in relation to a camera. Other data may also be included such as GPS data, device identification information, and the like." ('910 patent (EX1001), 6:31-41 (emphasis added).) This is inconsistent with his testimony where he agreed with the specification's description of "biometric data." (EX1059, 50:11-14.)

229. Dr. Turk's third reason is because Derakhshani's landmarks are not a mapping of facial features since it does not capture images of the user's face. (EX2061, ¶¶397-398.) As I discussed above, Derakhshani does capture images of faces, the images may include facial feature landmarks such as "an eye, nose, mouth, and ear." (EX1005, 17:16-19.)

7. Derakhshani and Tanii teach "reflection of the illumination from a face of the user" ('910 Patent, Claim 18), "face of the user is held steady" ('910 Patent, Claim 19), and "user's head is the user's face" ('910 Patent, Claim 24)

230. Dr. Turk's only argument why Derakhshani-Tanii does not teach claims 18, 19, and 24 is because Derakhshani captures images of the eyes and not the face. (EX2061, ¶370.) As I discussed above, Dr. Turk is wrong.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

B. Zhang And Tanii (Ground 2A)

1. Zhang and Tanii teach "comparing the first data to the second data to determine whether expected differences exist" ('910 Patent, Elements 1[d5] and 10[f]) and "comparing one or more aspects of the user's head … to determine whether expected differences, between the first image and the second image, exist" ('910 Patent, Element 22[d])

a. Zhang

231. As I explained in ¶¶59-64, Zhang teaches biometric authentication that involves distinguishing live faces from flat surfaces.

232. Dr. Turk acknowledges that Zhang teaches determining differences, but argues that it is not a determination of "expected differences" or "expected distortion." (EX2061, ¶402.) Dr. Turk reasons that color/intensity differences do not relate to expected distortion. (EX2061, ¶¶403-404.) This is false. A POSITA would have understood that color/intensity differences in images indicate depth, thereby captured images having color/intensity differences indicate a three-dimensional object such as a face exhibiting expected distortion. Zhang explains this: "[i]f the image difference meets the threshold value, then the face in the first and second images is determined to be a live face (act 314)." (EX1006, [0034].) Derakhshani explains this concept: "the intensity or color scale can be zoomed in to reveal more subtle changes in the level of the detected image signal. If the captured images are of a ***live eye, it is expected that the range of color or intensity values detected will continue to vary continuously***. In contrast, a spoofed image

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

(e.g., a digital photograph presented to the sensor) may exhibit large discontinuous jumps corresponding to halftones."). (EX1005, 19:30-37 (emphasis added).) Additionally, Dr. Turk's focus on a single pixel in each image is irrelevant because Zhang still determines the differences.

233. I further disagree with Dr. Turk's arguments that the color/intensity differences identified in Zhang does not relate to expected distortion that is based on movement of the camera or the user, and that Zhang makes no prediction of how an image would change due to this movement. (EX2061, ¶403.) This argument mischaracterizes Zhang's teachings and the Petition's Zhang-Tanii combination. Zhang discloses accessing "multiple images including a face presented by a user" and determining whether the face is a "3-dimensional structure or a flat surface" by analyzing differences between images. (EX1006, Abstract, [0026].) Zhang explains that when two images of a flat surface are compared, they can be aligned using a homography matrix such that "the warped image will be similar to the second image" with a "difference between the second image and the warped image [that] is small." (EX1006, [0025].) In contrast, when the images depict a three-dimensional face, "the difference between the second image and the warped image is larger" because the homography matrix cannot align different views of a 3D object. (EX1006, [0025].) The color/intensity differences that Zhang measures are the manifestation of this misalignment, which

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

in turn results from the three-dimensional geometry of the face creating perspective-dependent variations across the images. A POSITA would have understood that these color/intensity differences arise because different parts of the three-dimensional face are at different distances from the camera, causing them to appear differently across viewpoints in ways that a planar transformation cannot correct.

234. Moreover, the Zhang-Tanii combination does not rely on Zhang alone to teach analyzing expected distortion resulting from camera or user movement. Rather, as I explained in my opening declaration, the Zhang-Tanii combination relies on Tanii for teaching that three-dimensional faces exhibit predictable, distance-dependent perspective distortion. (EX1003, ¶¶1012-1022.) Tanii discloses "where the main object 9 and the cellular phone 1 are close together … an unnatural image results in which the perspective is exaggerated," and explains the technical reason: "the difference between the angle θ1 … and the angle θ2 … decreases as the lens unit 211 comes closer to the main object 9." (EX1007, [0047]-[0049].) Tanii further discloses that "the warp or distortion increases as the angle θ1 … increases" and that "where the main object is close to the imaging portion 2 and the image of the main object occupies a large percentage of the entire image, the exaggeration warp of the image becomes substantial." (EX1007, [0049].) These teachings provide the prediction Dr. Turk claims is missing: as the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

camera moves closer to a three-dimensional face, the perspective distortion increases in a predictable manner, with "peripheral areas of the main object appear[ing] reduced in size relative to the center area." (EX1007, [0056].) A POSITA would have understood that Zhang's framework would be enhanced by Tanii's teachings that the type and magnitude of differences expected from a three-dimensional face vary predictably with distance due to perspective distortion. The combination teaches comparing observed differences (as Zhang does via color/intensity analysis) to expected differences (as Tanii characterizes via distance-induced distortion) to determine three-dimensionality for authentication. Dr. Turk's argument that the color/intensity differences do not relate to expected distortion based on movement fails to account for Tanii's contributions to the combination.

235. Dr. Turk alleges that Jumio confuses Zhang's "feature points" and "selected location" to the "first data." (EX2061, ¶405.) As I explained in my opening declaration, the extracted feature points are data that characterize an individual's unique physical characteristics. (EX1003, ¶¶1041-1042.) The feature point location characterizes the location of the individual's unique physical characteristics. In other words, Zhang's feature point location is part of the feature point data, thereby showing that the mapping is not inconsistent.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

236. And Dr. Turk's arguments that the Petition cites to different passages in Zhang is flawed. (EX2061, 192 fn. 5.) Element 1[d2] and element 1[d5] recite different limitations. Element 1[d2] is directed to processing to create first data, while element 1[d5] is directed to comparing the first data to the second data to determine whether expected differences exist. Both sets of passages, [0027] (relied on for element 1[d2]) and [0031]-[0032] (relied on for element 1[d5]), discuss feature point data.

237. Further, Dr. Turk incorrectly tries to differentiate between feature point data used to generate the homography matrix H and feature point data used in the comparison of images. (EX2061, ¶¶405-409.) However, as Zhang explains, "feature points" are parts of the face, "such as a corner of an eye, a corner of a mouth, a tip of a nose, and so forth" that are extracted from the first and second images. (EX1006, [0027].) The extracted feature points are matched across the first and second images and used to generate the homography matrix H. (EX1006, [0028]-[0029].) Zhang discloses "[t]his matching of the feature points across the first and second images refers to identifying the locations of the same feature points in each of the two images. For example, the locations of the same corner of the same eye in the face in the first and second images are identified." (EX1006, [0028].) Then, when comparing the images, the image difference computes the "difference in color or intensity values at the same pixels (locations) in the two

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

images." (EX1006, [0031].) Zhang uses the same term "locations" when describing the generation of the homography matrix H and also when describing the difference in color/intensity values because it is referring to the same extracted feature points. Using Zhang's example of the corner of the eye as the "location," it would make sense for the comparison between the two images to be a comparison of the same feature point, such as the same corner of the same eye because otherwise a difference in color/intensity value may be due to different features points having different colors or intensity values due to, e.g., the user moving. For example, the color of the corner of an eye is likely to be different than the color of the corner of a mouth, thereby a comparison of the same pixel position having two different feature points (as Dr. Turk suggests) would not be indicative of whether the face is a 3D structure or a flat surface. Thus, Zhang's feature points are used for both generating the homography matrix H and the comparison of the images (part of authenticating the user).

238. Dr. Turk alleges that a three-dimensional image would not align with Zhang's approach and use of a homography matrix, and then concludes that Zhang does not predict how a three-dimensional face would distort across different views. (EX2061, ¶68.) However, Zhang's homography-based approach recognizes the fact that a three-dimensional face produces characteristic differences across views that are not eliminated by planar transformations. Zhang teaches that "two views

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

of a flat (planar) surface are related based on a homography matrix H" such that warping one image using the homography matrix H produces an image similar to the second image. (EX1006, [0025].) If there is a large difference between the first and second images, Zhang determines that "the first and second images include a face that is a live face rather than a flat surface." (EX1006, [0025].) A large difference is due to the images being of a three-dimensional face because different parts of the face are at different depths. The three-dimensional face undergoing different perspective transformations across viewpoints prevents the three-dimensional image from not being well-aligned using the homography matrix, as Turk alleges. (EX1006, [0030]-[0031]; EX2061, ¶68.) And a small difference indicates a flat surface (expected for a 2D photograph). (EX1006, [0025], [0033]-[0034].) A POSITA would have understood that Zhang's approach involves three-dimensional objects (including faces) that exhibit view-dependent distortions that are distinguishable from two-dimensional objects.

239. Moreover, as I explained above and in my opening declaration, the Petition combines Zhang with Tanii's teaching of the distance-induced distortion of a face when the camera comes closer to the face. (EX1003, ¶¶1012-1022.) Tanii teaches that perspective distortion of three-dimensional faces varies predictably with camera-to-subject distance, with closer distances producing more pronounced "exaggeration warp" where "peripheral areas of the main object appear

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

reduced in size relative to the center area." (EX1007, [0047]-[0049], [0056].) The Zhang-Tanii combination teaches a POSITA that (1) three-dimensional faces produce characteristic differences across images captured at different distances (Zhang's principle) and (2) these differences manifest as distance-induced perspective distortion (Tanii's teaching). Together, the Zhang-Tanii combination teaches determining whether expected distortion exists in the captured images, which is what the '910 claims require.

b. Tanii

240. Dr. Turk makes arguments about Tanii for the Zhang-Tanii combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii combination. (EX2061, ¶¶410-411, 423.) I explained why Dr. Turk's arguments are incorrect above.

2. A POSITA would have been motivated to combine Zhang and Tanii

241. Dr. Turk argues that there is no motivation to combine Zhang and Tanii because Tanii destroys Zhang's "basic objective" that requires distortions to operate. (EX2061, ¶¶436-437, 440.) Dr. Turk's arguments are however premised on his misunderstanding of the Zhang-Tanii combination. The teachings of Tanii that are part of the Zhang-Tanii combination are its teachings of distance-induced distortions in images of a user's face. (EX1003, ¶¶1012-1022.) Dr. Turk

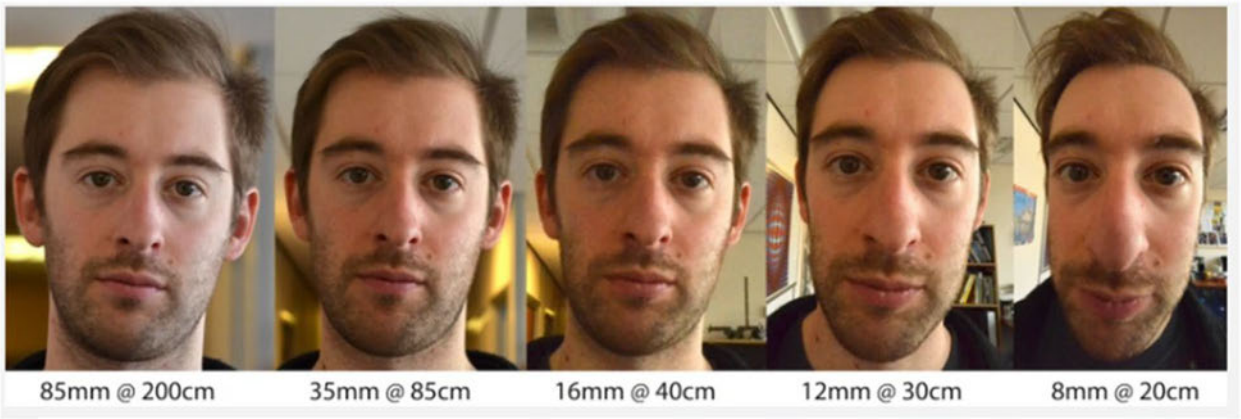
Jumio EX1060
Jumio v. FaceTec IPR2025-00109

erroneously bases his arguments on Tanii's teachings of image correction, which are not part of the Zhang-Tanii combination. (EX2061, ¶¶439, 446-447, 453.)

242. Dr. Turk also criticizes Tanii's figures for allegedly being inconsistent with its disclosures. (EX2061, ¶450.) Specifically, Dr. Turk alleges that the ears in the warped image (Figure 9) would not "miraculously appear after Tanii's purported distortion correction." (EX2061, ¶450.) However, Dr. Turk's argument shows that he misunderstands Tanii. A POSITA would have understood that Tanii's figures are simplified line drawings that portray abstract representations of the distance-induced distortions caused by the optical effects. (EX1007, [0048], [0049], [0056].) An actual captured image, on the other hand, has more than hundreds of pixels. Thus, Tanii's figures do not depict the level of detail of a captured image. Tanii does not cause ears to "miraculously appear," as Dr. Turk asserts. Instead, the magnification is changing based on the distance from the image center. (EX1003, ¶¶46-51.) Given that Tanii's figures are simplified and abstract depictions, a POSITA would have understood that the ears are present in the actual captured image corresponding to Figure 4B, but because of the magnification, the ears are difficult to see. For example, the rightmost image (below) has ears that are present, but is difficult to see with the human eye due the distance-induced distortion. In those photographs, the user's ears appear increasingly distorted (reduced in lateral size compared with other facial features)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

as the distance between the sensor and the user is reduced. Also consistent with Tanii's figures, the user's nose exhibits distortion where it is relatively enlarged as this distance is reduced.



3. Zhang and Tanii teach "verifying the presence of one or more features[/ears] … in the at least one first image, and verifying the absence or reduced visibility of the one or more features [/ears] … in the at least one second image" ('910 Patent, Claims 3 and 12)

243. Dr. Turk makes arguments about Tanii for the Zhang-Tanii combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii combination. (EX2061, ¶¶414-415.) As I explained above, Tanii teaches claims 3 and 12.

4. Zhang and Tanii teach "biometric data" ('910 Patent, Claims 7 and 16)

244. Dr. Turk argues that Zhang does not teach claims 7 and 16 because its feature points are not "biometric data" used for authentication. (EX2061, ¶¶416-418.) According to Dr. Turk, Zhang's feature points are used only to generate a

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

homography matrix H. (EX2061, ¶416.) This is the same argument that Dr. Turk made for elements 1[d5], 10[f], and 22[d]. I explained why Dr. Turk's arguments are incorrect above. Further, Zhang's facial features meet the description in the '471 specification of "biometric data" (similar to the description in the '910 specification), which Dr. Turk agrees with. (EX1059, 50:11-14.) Zhang states that "feature points are feature points of the face," "such as a corner of an eye, a corner of a mouth, a top of a nose, and so forth." (EX1006, [0027].) Thus, the comparison involves biometric data.

C. Zhang, Tanii, And Hoyos (Ground 2C)

245. Ground 2C challenges claim 18, which requires "reflection of the illumination from a face of the user." Dr. Turk makes arguments about Hoyos for the Zhang-Tanii-Hoyos combination that are, for the most part, the same as the arguments he made for the Derakhshani-Tanii-Hoyos combination. (EX2061, ¶¶419-422.) Dr. Turk also argues that Hoyos detects a reflection off the user's eye, not the user's face. (EX2061, ¶¶420-421.) Dr. Turk conflates Hoyos disclosures. Hoyos discloses at least two embodiments: the [0018] embodiment and the [0019] embodiment. In the [0019] embodiment, both "the first reflection [of the first image] and the second reflection [of the second image] were formed by a curved surface consistent with a human eye." (EX1010, [0019].) In the [0018] embodiment, only the "second image includes a representation of the first image

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

on the computer screen reflected by a curved surface consistent with a human eye." (EX1010, [0018].) In other words, while both images in the [0019] embodiment are a reflection off of the user's eye, the [0018] embodiment discloses only the second image being a reflection off of the user's eye. As a result, the first image in the [0018] embodiment may be a reflection off of the user's face. Dr. Turk testified during his deposition that the reflection off the eye in the [0018] embodiment is for the second image only. (EX1059, 123:6-10.) Dr. Turk's other arguments are irrelevant or incorrect. Dr. Turk's argument that the surface of an eye for reflection is not the same as the surface of a face is irrelevant because claim 18 is not limited to a specific type of surface for the reflection. (EX2061, ¶421.) And Dr. Turk's assertion that Hoyos states it detects reflections from the eye only is incorrect. (EX2061, ¶421.) Instead, Hoyos states the "system expects to see a change ***primarily on the cornea***, where a sharp image of the screen is reflected." (EX1010, [0035] (emphasis added).) Being "primarily on the cornea" means the reflection is mostly, but not entirely, on the cornea. Dr. Turk agreed. (EX1059, 123:11-124:8.) This is consistent with the reflection in the [0018] embodiment being off of the user's face.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

A. Tanii Is Analogous Art

1. Tanii is from the same field of endeavor

246. Dr. Turk argues that Tanii is not in the same field of endeavor as the ’910 patent because the ’910 uses distortion and Tanii tries to remove it as part of its image correction process. (EX2061, ¶¶424-425.) But Dr. Turk’s argument that Tanii assumes distortion exists and corrects for it is misplaced. The Derakhshani-Tanii combination and the Zhang-Tanii combination rely on Tanii for its teachings of distance-induced distortion, not on its teachings of an image correction process.

247. Further, a POSITA would not have understood Tanii’s teachings to be so narrow as Dr. Turk argues. Instead, a POSITA would have understood that one of Tanii’s teachings is related to identifying and accounting for distance-induced distortions of three-dimensional faces in images captured using a camera.

248. Thus, both the ’910 patent claims and Tanii relate to distance-induced distortion. That is, Tanii is directed to the same function (accounting of distance-induced distortions) and structure (camera), and as a result is in the same field of endeavor as the ’910 patent.

2. Tanii is reasonably pertinent to the problems faced by the inventors of the ’910 Patent

249. Dr. Turk acknowledges that one of the problems the ’910 inventors were faced with is “this long-understood issue” of a three-dimensional face exhibiting distortions when the camera and face are close together during image

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

capture. (EX2061, ¶430.) Dr. Turk does not dispute that Tanii is reasonably pertinent to this distance-induced distortion problem, but instead argues that Tanii's teachings would not be helpful because the problem was known. (EX2061, ¶430.) But whether Tanii's teachings were known does not determine whether Tanii is reasonably pertinent.

250. I note that the Petition stated that "[t]he '910 Patent describes . . . methods to distinguish a three-dimensional face from a two-dimensional picture of a face." (Petition, 4.) The Petition did not admit that distinguishing a three-dimensional face from a two-dimensional picture of a face was a problem faced by the '910 patent, as Dr. Turk alleges. (EX2061, ¶429.)

251. Dr. Turk argues that the '910 patent solves the spoofing problem and Tanii does not. (EX2061, ¶¶431-432.) But spoofing is not the only alleged problem faced by the '910 inventors. As discussed above, Tanii solves the '910 problem related to distance-induced distortions.

252. Dr. Turk criticizes Tanii because allegedly "[t]he '910 patent uses the distortion that Tanii tries to remove." (EX2061, ¶425.) He argues that "Tanii works against the challenged claims, by removing the very thing that enabled the patent's inventor to solve the spoofing problem." (EX2061, ¶434.) I disagree. A POSITA would have understood that Tanii's teachings of distance-induced distortion is suitable for use for user authentication.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

253. And Dr. Turk's argument about purpose is flawed. (EX2061, ¶435.) As described in the context of the relevant art is discussed, [0006] states "[i]f the image of the face of a person or the like is captured from a point that is relatively nearby, an image is obtained in which the perspective appears exaggerated. . . . When a camera comes close to a three-dimensional object that has depth, the peripheral areas of the object, which are farther away from the camera than the front area, appear to be closer to the front area than they really are, resulting in a warp that causes the image to appear as if the perspective were exaggerated (the type of warp in which the perspective becomes exaggerated, as in this case, will hereinafter be termed 'exaggeration warp', as distinguished from distortion)." (EX1007, [0006].) Thus, Tanii accounts for warp that causes the image to appear as exaggerated.

VII. OBJECTIVE EVIDENCE OF NONOBVIOUSNESS

A. The Challenged Claims Are Not Coextensive With FaceTec's ZoOm Product

254. I understand that objective evidence of nonobviousness must have a nexus to the claimed invention, and that such a nexus may be presumed only where the evidence is tied to a specific product. Such a product must both embody the claim and be coextensive with it. I understand that a product is not coextensive with the claims if it includes critical unclaimed features, particularly those that are claimed by a different patent or patent application.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

255. I understand that the challenged independent claims are directed to determining whether a user or user's face is three-dimensional.

256. For example, independent claim 1 of the '471 patent recites "[a] system for *authenticating three-dimensionality of a user* via a user's camera equipped computing device, the computing device comprising . . . one or more memories configured to store machine readable instructions . . . which when executed by the processor, cause the computing device to: capturing at least one first image of the user . . . at a first location which is a first distance from the user; processing the at least one first image or a portion to create first data; . . . capturing at least one second image of the user . . . at the second distance from the user, the second distance being different than the first distance; processing the at least one second image or a portion thereof to create second data; *comparing the first data to the second data to determine whether expected differences exist between the first data and the second data which indicated three-dimensionality of the user; authenticating the user when differences between the first data and the second data have expected distortion resulting from movement of the camera from the first location to the second* . . . ." (EX1001 of IPR2025-00106, Claim 1 (emphasis added).)

257. Independent claim 1 of the '938 patent recites "[a] non-transient computer readable medium containing non-transitory machine executable code

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

configured to determine if the three-dimensional shape is consistent with that of a human face, the non-transitory machine executable code configured to: receive or derive first biometric data from at least one first image of a user taken with a computing device camera located at a first distance from the user; receive or derive second biometric data from at least one second image of the user taken with the computing device camera located at a second distance from the user, the second distance being different than the first distance; compare the first biometric data with second biometric data for expected differences that result from characteristics of a human face and the at least one first image and the at least one second image being captured at different distances from the user; *determine that the three-dimensional shape is not exhibited when the second biometric data does not have expected differences compared to the first biometric data . . . .*" (EX1001 of IPR2025-00108, Claim 1 (emphasis added).)

258. Independent claim 1 of the '606 patent recites "[a] method for *verifying three-dimensionality of a user's face* using images of the user's face . . . the method comprising: capturing at least one first image of the user taken with the camera of the computing device at a first distance from the user; processing the at least one first image to obtain first biometric data from the at least one first image; capturing at least one second image of the user taken with the camera of the computing device at a second distance from the user, the second distance being

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

different than the first distance; processing the at least one second image to obtain second biometric data based on the at least one second image; . . . comparing the first biometric data to second biometric data to determine whether differences between the at least one first image and the at least one second image match expected differences resulting from movement of the camera or the user which changed the distance between the user and camera from the first distance to the second distance; *determining that the user's face is three-dimensional when: . . . the second biometric data has the expected differences as compared to the first biometric data resulting from the change in distance between the user and the camera when capturing the at least one first image and the at least one second image*." (EX1001 of IPR2025-00107, Claim 1 (emphasis added).)

259. Independent claim 1 of the '910 patent recites "[a] computing device for *verifying three-dimensionality of a user* via a user's camera equipped computing device, the computing device comprising: . . . one or more memories configured to store machine readable instructions that are stored on the memory of the authentication server which when executed by the processor, cause the computing device to: capturing at least one first image of the user taken with the camera of the computing device at a first location which is a first distance from the user; processing the at least one first image or a portion thereof to create first data; capturing at least one second image of the user taken with the camera of the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

computing device is at a second distance from the user, the second distance being different than the first distance, the capturing at least one second image of the user occurring after movement of the camera or the user to establish the camera at the second distance from the user; processing the at least one second image or a portion thereof to create second data; *comparing the first data to the second data to determine whether expected differences exist between the first data and the second data which indicates three-dimensionality of the user; verifying the images of the user exhibit three dimensional traits when the expected differences exist between the first data and the second data . . . .*" (EX1001 of IPR2025-00109, Claim 1 (emphasis added).)

260. I understand that the remaining independent claims are all also directed to determining whether a user or user's face is three-dimensional.

261. As discussed above, I understand that the analysis for secondary considerations considers whether a product embodies the challenged claims. I understand that FaceTec's objective evidence is based on the ZoOm product, with the name of the product changing after the pandemic to FaceTec SDKs. (EX2061, ¶456, EX1059, 82:10-16; EX1082, 29:17-31:5; EX2062.) For purposes of this declaration, I will refer to the product as "ZoOm." I understand that Dr. Turk's views FaceTec's ZoOm as embodying the independent claims of the Challenged Patents, including the four independent claims set forth above. (EX2061, ¶456;

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

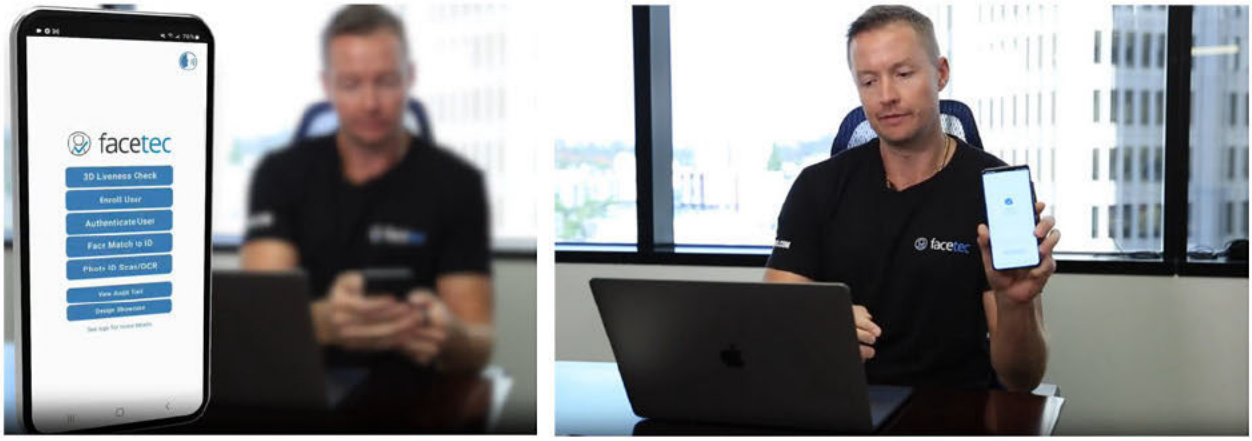
EX1059, 82:10-16.) Specifically, Dr. Turk analyzed only claims 1 and 10 of the ’471 patent, claims 1, 10, and 19 of the ’606 patent, claims 1, 8, 17, and 20 of the ’938 patent, and claims 1, 10, and 22 of the ’910 patent with respect to his nexus analysis. (*See* EX2061, Appendices A-H.) I understand that Dr. Turk did not analyze any other claims of the Challenged Patents.

262. According to FaceTec and Dr. Turk, the software (aka ZoOm) runs on a computing device having a camera, screen, processor, and memory. (POR, 6; EX2061, ¶456; EX1059, 82:10-16.) When ZoOm is initiated by a user, the software performs a number of actions to ultimately provide a “high confidence decision” that “a real live human being [is] in front of the camera and not some sort of spoof or attack.” (EX1082, 75:6-21; EX2059, ¶49 (liveness check is “multilayered”); [redacted] (EX1082, 43:24-44:18; EX2069, 0:20-0:45.) One of the checks at the server determines whether the user is three-dimensional. (EX1082, 99:15-100:3, 149:3-16; EX1059, 78:23-79:20, 114:2-115:4.) Specifically, the software confirms that a user’s face is three-dimensional by capturing at least one facial image of the user at a first distance from a camera, capturing at least one facial image of the user at a second distance from the camera, and comparing data from the images for

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

expected differences, such as the presence of appropriate perspective distortion. (POR, 6; EX2059, ¶¶9-15, 16-21, 23-24; EX2061, ¶456.)

263. I understand FaceTec has submitted a number of exhibits explaining the operation of ZoOm. One exhibit is a video (EX2069) that shows Josh Rose, FaceTec's CTO, having his "liveness" confirmed while using a smartphone. As shown below, the check begins when a user, Mr. Rose, clicks the "3D Liveness Check" button on the user interface and continues until the interface notifies him that liveness is confirmed.



(EX2069, 0:28 (initiating a 3D liveness session), 0:43 (noting that liveness is confirmed).) Dr. Turk testified that he was referring to the portion of the video from about 0:20 seconds to about 0:45 seconds for his secondary considerations analysis. (EX1059, 84:19-85:2.)

264. [redacted] Dr. Turk relies upon EX2067 [redacted] (EX1059, 99:11-17.) Dr. Turk testified [redacted]

[redacted] (EX1059, 111:3-25.) EX2067 provides [redacted]

265. [redacted] Even though FaceTec utilizes the term "3D liveness," the decisioning process is not limited solely to determining liveness based on three-dimensionality. [redacted] Thus, while he testified that the "core functionality" of FaceTec's software is to determine three-dimensionality (EX2061, ¶456), he understood that "3D liveness detection can include several other ways of providing evidence to come to that conclusion." (EX1059, 79:8-13.) I understand that Mr. Rose similarly testified that "the 3D liveness checks are multilayered" and [redacted]

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

[redacted] (EX2059, ¶49; [redacted]) [redacted]

[redacted] (EX1082, 99:15-100:3, 149:3-16.)

266. Indeed, my understanding is that FaceTec's ZoOm has at least the following additional "layers" or "features": 1) control of the camera, 2) a user interface (UI), 3) encryption of data before being sent to a server, 4) liveness checks, and 5) removal of the liveness data. At least each of these features is critical, according to FaceTec, to confirming whether the user, such as Mr. Rose in the video, is confirmed to be a live person using FaceTec's ZoOm. Also, at least these features are not claimed by the independent claims of the Challenged Patents, because the independent claims are directed to determining the three-dimensionality of a user, as I explained above. Accordingly, in my opinion, FaceTec's ZoOm is not coextensive with the independent claims of the Challenged Patents.

1. Control of the camera

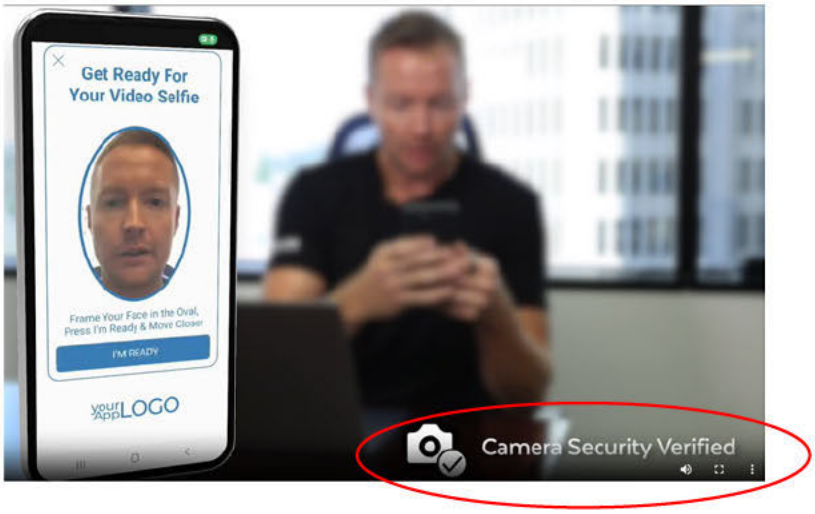
267. I understand that FaceTec's ZoOm controls the camera and verifies the feed coming from it. I understand that PWC conducted an interview with FaceTec management, including its CEO Kevin Alan Tussy and its chief strategist, Jay Meier, in 2021. I understand that the summary of the interview is available on FaceTec's website. According to the summary, FaceTec management explained

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

that their "technology works . . . first by ensuring camera feed integrity." (EX1067, 2.)

268. Mr. Meier in a more recent podcast interview provided one reason why this feature is important. He said that FaceTec's Software Development Kit or SDK "take[s] control of the . . . camera and the data feed from it" to help defend against deepfake injection attacks. (EX1073, 11:1-4; EX1086/EX1072.) According to Dr. Turk, deepfakes "are typically videos – it could be audio or other things – that are made intentionally to pretend to be someone but are actually not." (EX1059, 89:20-90:4.) In its 2021 interview with PWC, FaceTec explained these kinds of threats are "the most difficult to defend" and are "the #1 threat to be concerned about in the future." (EX1067, 3.)

269. As shown by Mr. Rose in the video discussed above, after clicking the "3D Liveness Check" button and the ovals appearing, the camera security is verified.



(EX2069, 0:31.) Mr. Rose testified, after seeing this video at his deposition, that [redacted] (EX1082, 38:15-39:6.) In his declaration, Mr. Rose testified about [redacted] (EX2059, ¶46.) [redacted] (EX1082, 43:12-45:3.) Additionally, according to a FaceTec quarterly Liveness Detection report, the software "determin[es] that the app is not running on an emulator, a virtual camera is not being used, and a camera hardware adapter bypass is not being attempted." (EX1065, 3; *see also* EX1064, 3.) [redacted] (EX1082, 41:15-22.)

270. Controlling the camera is important, according to FaceTec, for another reason: obtaining the best feed possible for performing the liveness check.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

FaceTec's ZoOm controls the camera parameters to obtain "the highest quality video data possible from the device at the maximum camera resolution and frame-rate." (EX1067, 2.)

271. Control of the camera and feed are required for ZoOm, according to FaceTec. On its Frequently Asked Questions page, FaceTec notes that the "SDKs are a *required* integration component" that can, among other abilities, "selec[t] the optimal resolution and processing parameters depending on the device, camera properties, platform, screensize, and estimated device performance profile." (EX1086/EX1061, 4 (emphasis added).) Further, "failing to test for real-time data collection by the camera and related data flows could effectively enable the most scalable biometric system attacks to be successful against FRVT-PAD-tested vendors." (EX1064, 3.)

272. The independent claims of the Challenged Patents are directed to authenticating or verifying three-dimensionality by looking at expected differences between images, as I have discussed above. They do not claim controlling the camera and verifying feed integrity. In fact, FaceTec has at least one other patent application, US 17/939,867, directed to this unclaimed feature. The application is entitled "Facial recognition and/or authentication system with monitored and/or controlled camera cycling." (EX1074.) [redacted] The

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

application claims "[a] method for detecting video feed injection during an authentication session." (EX1074, claim 1.) The application specification notes that the method for detecting video feed injection is for a session that "may be any type or authentication . . . such as . . . three-dimensionality detection of the user's face." (EX1074, [0268].)

273. Mr. Rose testified that "the spoof and security landscape definitely evolves." (EX1082, 113:6-16.) Dr. Turk explained that "any company would want to be evolving their system to meet the latest kinds of attacks that they might be facing." (EX1059, 171:14-21.) To that end, Mr. Rose explained that FaceTec came out with "Level 4 and Level 5 on the Spoof Bounty program before there were any types of standards for testing that as well." (EX1082, 113:6-16.) I understand that Level 5 refers to "successfully tak[ing] over the camera feed & inject previously captured frames that result in the FaceTec Server responding with Liveness Success." (EX2046, 1.) Furthermore, Mr. Rose explained that 3D liveness checks are "multi-layered" [redacted] (EX2059, ¶46; [redacted]) I understand that Mr. Rose refers to this layer as "incremental," but this testimony is not supported by FaceTec's statements to the public as quoted above (video injection and virtual cameras "most difficult to defend") and FaceTec's Spoof Bounty program. FaceTec certainly thought controlling the camera and feed to be a

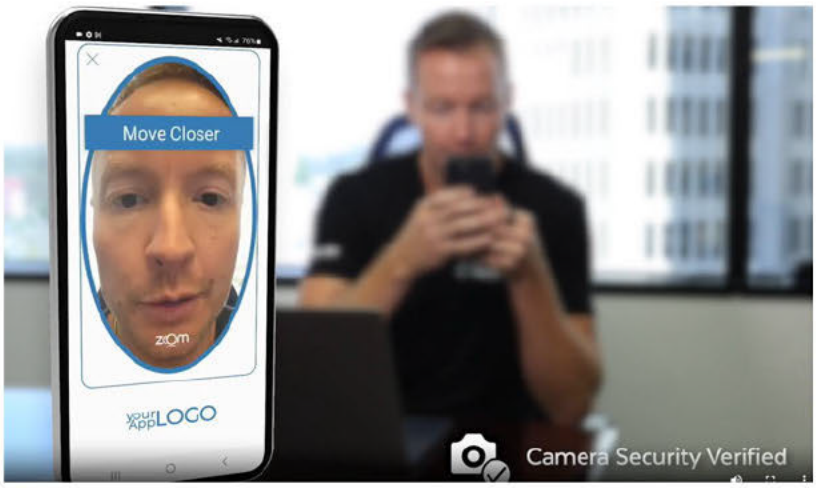
critical feature in the evolving spoof and security landscape, so much so that they created a Level 5 threat in its Spoof Bounty Program. Indeed, FaceTec defines the goal of the Spoof Bounty Program as follows: "The goal of the Bounty Program is to uncover any potentially unknown vulnerabilities in the Liveness AI and security layers so that any issues can be patched and the anti-spoofing capabilities and overall platform security elevated even further." (EX2046, 1.) Mr. Rose makes similar statements in his declaration. (EX2059, ¶¶29-31.) Therefore, the control of the camera and feed is a critical feature, according to FaceTec, to prevent against, for example, injection attacks. It is also feature that is unclaimed by the independent claims reviewed by Dr. Turk with respect to nexus.

2. UI

274. I also understand that FaceTec's ZoOm has a UI involving two ovals. FaceTec notes that their 3D liveness detection technology includes use of a UI with a small oval to capture a far image and a larger oval for capturing a closer image. (POR, 8; [redacted].) The UI appears when the user presses the "3D Liveness Check" button in the video with Mr. Rose.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

(EX2069, 0:31 (smaller oval), 0:34 (large oval); [redacted]) The user "aligns or places their face [in the ovals] during the capture" of the images. (EX2059, ¶14.) When the user moves between the two ovals, Mr. Meier explains that about 100 different video frames are captured. (EX1073, 3:3-4:17.)



(EX2069, 0:34.) Dr. Turk understood that the video of Mr. Rose shows ZoOm encompassing "the generation of ovals." (EX1059, 87:8-16.) [redacted] (EX1082, 109:10-24, 20:6-13.)

275. [redacted] Mr. Rose testified that "the user action that we trigger by engineering the UI in the way that we have has kind of the stated goal of having the user change the distance between themselves and the camera. So it's a fundamental aspect of what we do." (EX1082, 134:17-23.) Mr. Meier provided further explication why, according to FaceTec, the UI is important, namely, for controlling the capture of images through the use of ovals. According to Mr. Meier, the use of the two ovals "intentionally create[es] variability in the depth of field . . . in a way that is consistent across faces." (EX1073, 3:9-11.) The two ovals are "the *only* way [FaceTec] could figure out how to capture 100 different video frames in a very controlled environment." (EX1073, 6:17-21 (emphasis added).) Mr. Rose highlighted that the user interface is a "security feature" that "*requires* users to provide images of their face from at least two different distances." (EX2059, ¶46.) FaceTec also emphasized the ovals in arguing that its technology was praised. (POR, 8-9.)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

276. The independent claims of the Challenged Patents do not claim the UI with respect to the ovals. In fact, FaceTec has several other patents directed to this unclaimed feature. For example, FaceTec has at least two design patents directed to the user interface:

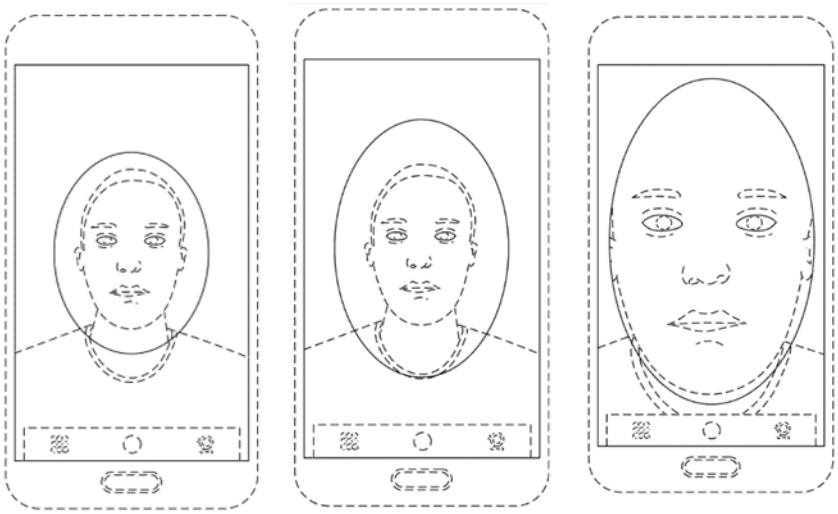


FIG. 1 FIG. 2 FIG. 3

(EX1075, FIGs. 1-3.)

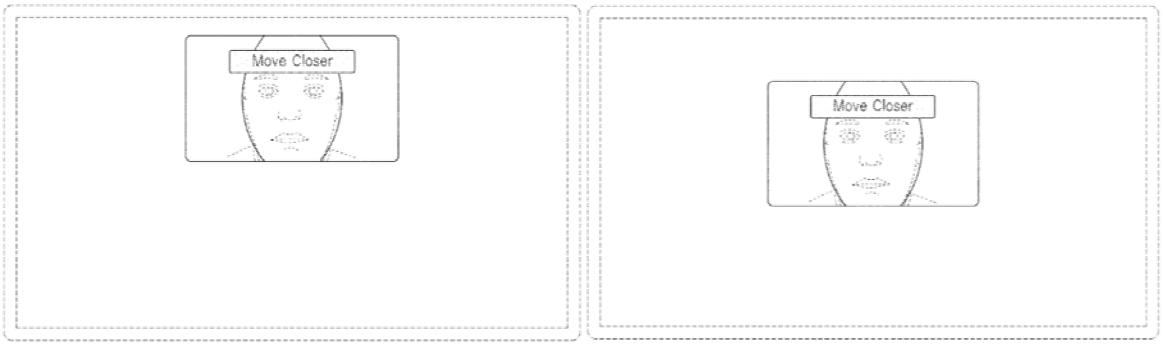


FIG. 1 FIG. 2

(EX1076, FIGs. 1-2.) I understand that, in a press release, FaceTec explained that at least the above first design patent USD 1074689 covers the FaceTec's "3D Liveness detection capabilities." (*See* EX1086/EX1066, 1.)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

277. FaceTec also has at least one utility patent, US 12,423,398, directed to this unclaimed feature. (EX1077.) The patent claims "[a] system for capturing images of a person's face and evaluating one or more characteristics of the person's face" wherein the system prompts the user with a "face-framing prompt" and captures and processes the images to determine, among other things, "one or more characteristics of the person's face captured in the images and/or verify that the camera captured images of a real person." (EX1077, Claim 1.) The patent further claims that the "face-framing prompts are an oval shaped guide on the screen within which the face of the person is aligned to capture the at least one first image and the at least one second image." (EX1077, Claim 11.)

278. Therefore, the UI with ovals is a critical feature, according to FaceTec, of ZoOm. It is also a feature that is unclaimed by the independent claims reviewed by Dr. Turk with respect to nexus.

279. I understand that the same evidence of secondary considerations cannot be presumed to be attributable to two different combinations of features. I understand that at least the '938 Challenged Patent lacks any claim directed to a UI with ovals, yet FaceTec presents the same evidence for all four proceedings. (*See* EX1001 of IPR2025-00108.) Even if FaceTec argued that some Challenged Patents may have a dependent claim directed to using ovals, the same cannot be

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

said of all four of the Challenged Patents, and the evidence provided cannot be independently attributable to patents claiming different features.

3. Data encryption

280. I understand that FaceTec's ZoOm encrypts the data acquired on the device side and securely sends it to the server during a liveness check, as shown in the image below, for decryption and processing. According to FaceTec, the captured images "are then processed on the device itself, producing a small file containing 3D FaceScan data. This 3D FaceScan file (~350kb) is then encrypted and sent to the server for analysis." (EX1067, 2; [redacted])

(EX2069, 0:37.) [redacted]

[redacted] Dr. Turk further confirms the encrypted data is then "evaluated on a server" as shown above. (EX1059, 86:22-87:7.)

281. According to FaceTec, encryption of the FaceScan data for decryption by the secure server is not only critical but also required for ZoOm. Mr. Rose testified that "[t]o ensure the security of the system as a whole, it's important that the data we process and analyze and inspect and pretty anything we do with the data hasn't been tampered with." (EX1082, 35:2-37:7.) He further testified that "once we create this 3D face scan, which is an encrypted piece of data—once that piece of data exists, we don't want a potential attacker to be able to decrypt that and be able to inspect or modify or tamper with that piece of data." (EX1082, 35:2-37:7.) FaceTec similarly notes that ZoOm cannot be used without server-side processing; otherwise "Video Injection, GenAI Attacks, and Root-Bypass Attacks would be very easy to perform." (EX1061, 3.) Mr. Meier, when discussing encryption of the data, confirms that "you have to use cryptography. . . . If you don't, you're going down in a hurry. . . Cryptography is essential." (EX1073, 16:17-19, 19:19-20.) Given that the FaceScan data from a user must be received by a server, it is critical, according to FaceTec, that the data be first encrypted lest

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

a bad actor accesses the data on its way to the server for decryption and processing. I understand FaceTec has created a Level 4 threat in its Spoof Bounty program given the importance of the threat. (EX2064, 1.) The threat is successfully decrypting and editing the contents of a 3D FaceScan. (EX2064, 1; EX1082, 35:2-37:7; [redacted] As I noted above, FaceTec defines the goal of the Spoof Bounty Program as follows: "The goal of the Bounty Program is to uncover any potentially unknown vulnerabilities in the Liveness AI and security layers so that any issues can be patched and the anti-spoofing capabilities and overall platform security elevated even further." (EX2046, 1.)

282. The independent claims of the Challenged Patents do not claim encrypting data for processing of data on a secure server.

283. In fact, FaceTec has at least one other patent application, US 17/966,639, directed to this unclaimed feature. The application claims "[a] node server for autonomously authenticating a user" comprising "an authentication SDK configured to perform authentication using data from the node database" and "an access module configured to permanently and irrevocably prevent access from external servers or devices not associated with a user account, after creation and activation of the node server." (EX1078, Claim 1.) The application specification notes "the application running on the authentication device 3612 creates a 3D

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

facemap which may be encrypted . . . The 3D facemap may comprise biometric data, image data, or any other type data regarding the user's face as derived from one or more images of the user. . . . In one embodiment, the image(s), facemap, . . . or 3D facescan . . . are encrypted using https encryption when being sen[t]to the server that compared to the newly captured image to the f[l]at image, such as the server SDK # 1. The one or more images captured during an authentication session may be processed to create biometric information . . . [which] can be used to verify the liveness of the user. . . . One type of biometric information referred to herein is a 3D facescan which is contains or includes biometric information that can be used to verify liveness and biometric information that is or can be used to generate a facemap." (EX1078, [0351].)

284. Therefore, the data encryption feature for server-side decryption and processing is a critical feature, according to FaceTec, of ZoOm to prevent, for example, middle-man attacks. It is a feature that is unclaimed by the independent claims reviewed by Dr. Turk with respect to nexus.

4. Liveness checks

285. I understand that FaceTec believes its challenged independent claims are coextensive with a check that looks for perspective distortion between images to determine three-dimensionality. For example, Dr. Turk notes that "FaceTec's software embodies and is coextensive with the challenged claims. Indeed, the core

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

functionality of FaceTec's software . . . is achieved by using a camera-equipped computing device having a camera, screen, processor, and memory to capture facial images of a user at two distinct and fairly consistent distances, process the images, and analyze data from images for expected differences in, e.g., perspective distortion." (EX2061, ¶456.)

286. However, determining liveness solely based on three-dimensionality has limits, as "true liveness detection requires wholistic observation of the characteristics of the face and the surrounding environment." (EX2036, 3.) In particular, as Dr. Turk explained, [redacted] (EX1059, 100:24-101:6.) Jay Meier confirms this, adding that "[i]f the only thing you measure is three-dimensionality, then those masks will work." (EX1073, 12:24-13:1.) Indeed, Dr. Turk explained that "any company would want to be evolving their system to meet the latest kinds of attacks that they might be facing." (EX1059, 171:14-21; *see also* EX1086/EX1071, 1 (noting that the Spoof Bounty Program is FaceTec's way of "hunt[ing] for ways to defeat Face[T]ec's 3D Liveness software in real-world conditions, demonstrating and improving upon its security in real-world conditions, not just in the lab[.]").) Hence, Mr. Rose has admitted that it takes a "multilayered" approach. (EX2059, ¶49.) Dr. Turk agreed, and added that "3D liveness detection can include several other ways of providing

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

evidence" besides determining whether a user is three-dimensional. (EX1059, 79:8-13.)

287. I understand that, as part of this multilayered approach, FaceTec's ZoOm performs liveness checks, in addition to the three-dimensionality check. (*See* EX2059, ¶49.) Mr. Rose testified that [redacted] (EX1082, 46:21-47:10.) Mr. [redacted] (EX1082, 79:10-24.) Mr. Rose was later asked [redacted] (EX1082, 71:15-72:22.) Mr. Rose testified that [redacted]

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

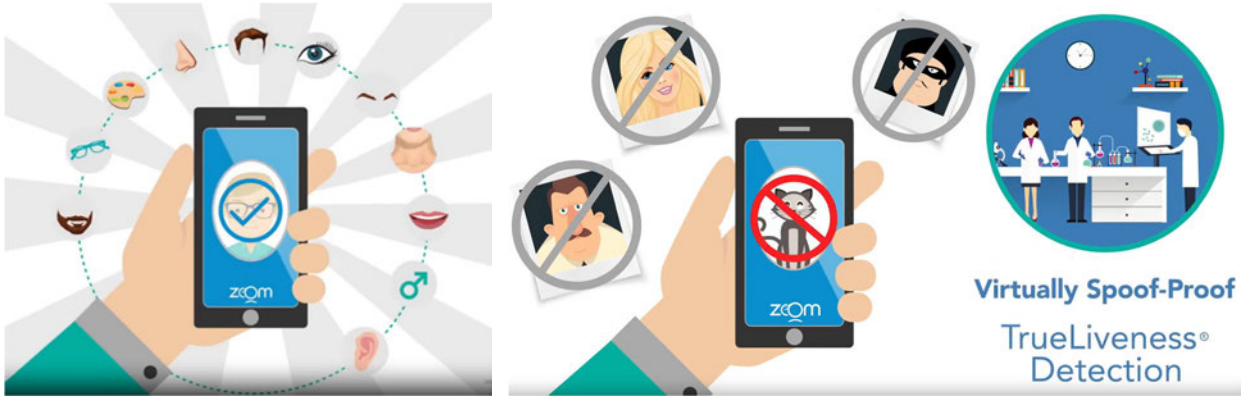
[redacted] (EX 1082, 85:22-86:4.) He explained near the end of his deposition that [redacted] (EX1082, 149:3-16.)

288. According to a FaceTec Case Study, "FaceTec's software measures perspective distortion to prove the user's face is 3D *and* measures several dozen uniquely human traits." (EX1063, 4 (emphasis added).) The product "provides passive liveness via involuntary human signal measured by deep neural networks, which validate up to 60 human traits." (EX1067, 3.) The system is able to look at involuntary cues FaceTec (Mr. Meier) calls "passive liveness characteristics." (EX1073, 13:1-22.) I understand that Mr. Rose testified that "'[p]assive liveness check' is a heavily loaded term with a lot of different definitions depending on who you talk to. It's not something we use on the product building or the technology side." (EX1082, 117:23-118:6.) I note that both Dr. Turk and Mr. Meier used that term as discussed immediately above and below. Regardless of the term, it is clear from the testimony of Dr. Turk and Mr. Rose that ZoOm performs liveness checks beyond checking for three-dimensionality.

289. Dr. Turk admits that "passive" liveness tests can provide an "additional bit of information to help . . . the liveness determination." (EX1059,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

117:10-19.) One example is looking at how the pupils dilate under different conditions and reactions in the user's eye focus. (EX1073, 13:7-13; EX1067, 3; [redacted].)



(EX2062, 0:41, 0:50.)

290. I understand that the screenshots of the ZoOm video above show that the ZoOm product looks at a number of different things beyond just three-dimensionality. (EX2062, 0:41, 0:50.) For example, the video mentions that ZoOm uses "artificial intelligence to measure your uniquely human traits." (EX2062, 0:50.) Mr. Rose also testified [redacted] (EX1082, 119:24-121:4.)

291. These liveness checks are a critical feature, according to FaceTec, of ZoOm and are performed "in addition to the three dimensionality" check. (EX1073, 13:20-22.) While certain approaches, such as a three-dimensionality check, "can be effective in discerning three-dimensional from two-dimensional

objects, there is one key issue: a 3D object is not necessarily a real person. Dolls, masks 3D prints and wax figures are all 3D, but not alive." (EX1063, 3.) The ability to detect "several dozen uniquely human traits" helps provide a more accurate determination that someone is alive. (EX1063, 4.) FaceTec claims that this feature is "why our liveness is so much better than anything out there" – because it confirms a user is three-dimensional which means "[p]resentation attacks are largely gone" and then it uses "70 or more involuntary human cues that no one knows what they are so they don't know what to attack" as part of these liveness checks. (EX1073, 14:9-14.)

292. FaceTec's own patents distinguish between authentication of three-dimensionality and authentication of "liveness." For example, US 11,693,938, one of the challenged patents, claims that "authentication is authentication of liveness, three-dimensionality, or both," explicitly separating the two types of authentications. (EX1001 of IPR2025-00108, claim 17.) The specification of the '938 patent further clarifies that the method may use "eye movement, blinking, and the like to determine the 'liveness' of the person being imaged," methods that are not included within the scope of the independent claims directed solely to a three-dimensionality check. (EX1001 of IPR2025-00108, 30:65-66.)

293. Other FaceTec patents/patent applications make a distinction between "liveness" and "three-dimensionality" checks. For example, US 12,130,900 claims

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

"capturing one or more images of the user's face with the camera during an authentication session, wherein the authentication session comprises one or more of the following: *identity analysis, facial three-dimensionality analysis, and liveness analysis*." (EX1079, Claim 8 (emphasis added).) US 12,130,900 also claims "[a] method for authenticating a user" involving "verify[ing] that the one or more images were captured from a three-dimensional face," but that "the authentication *further comprises liveness analysis*." (EX1079, Claims 1-2 (emphasis added).) These additional "liveness" checks are not included in the scope of the claims of the Challenged Patents. US 11,657,132 claims "[a] method for authenticating" wherein "authenticating comprises performing *liveness verification, three dimensionality verification, or both*." (EX1081, Claim 1 (emphasis added).)

294. Therefore, liveness checks are a critical feature, according to FaceTec, of ZoOm to prevent against attacks by, for example, 3D masked attackers. It is a feature that is unclaimed by the independent claims reviewed by Dr. Turk with respect to nexus.

5. Removal of liveness data

295. I understand that FaceTec's ZoOm generates a "3D FaceMap" upon confirming liveness. Mr. Rose testified that "[a] 3D face map is generated when the FaceTec server SDK evaluates a 3D face scan for 3D liveness and makes a

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

high-confidence decision that the scan that was performed came from a real human 3D face." (EX1082, 63:14-20.) He also testified that "[s]o when that all happens and there's a successful liveness determination, the software will output, in addition to the binary signal that – of whether there was success or not, will output a 3D face map, which is another encrypted piece of data that – essentially the use case for that is so that our customers can store the 3D face map. And, in general, we can call the storage of 3D face maps is enrolling or enrollment." (EX1082, 57:12-20.) Mr. Rose further explained: [redacted] (EX1082, 58:2-7.) According to FaceTec, "[o]nce liveness is confirmed on the server, the one-time use liveness data is deleted, leaving only 3D matching data, called a 3D FaceMap." (EX1067, 2.) FaceTec's own website defines a 3D Liveness Check as including the removal of liveness data, defining 3D Liveness Checks in their Technical Glossary as: "User Sessions produce FaceScans. FaceScans are evaluated by the FaceTec Server. FaceTec's highly secure and highly accurate ensemble of 100+ Algorithms and heuristics evaluate the FaceScan and if Liveness is proven, a result is given to you in your secure Server environment. *3D FaceMaps are provided as an output to the Liveness Check*

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

API." (EX1083, 1 (emphasis added).) This process of removing the liveness data eliminates the "honeypot" risk, that is, preventing theft and fraud of that liveness data. (*Id.*; EX1065, 6.) This is because "the data that is left over . . . is useless [because] [i]t can't be resubmitted. So . . . [e]ven if they lost everything or it was stolen . . . there would be no use for it." (EX1073, 9:21-10:2.) Mr. Rose similarly testified that [redacted] (EX1082, 58:18-59:1.)

296. Removal of the liveness data is a critical feature, according to FaceTec, of Zoom. FaceTec calls it a "*critical* selling feature in our view" with regards to preventing honeypot risk. (EX1073, 9:20-10:4 (emphasis added).) FaceTec also notes that "[w]ithout the sophistication of new 3D FaceMap technology," (which is the resulting file once the liveness data is removed), "the 2D face recognition solutions rarely saw any large-scale deployments . . . 3D FaceMaps contain as much as 100 times more data than a 2D photo . . . and are required to accurately . . . [verify] human 'liveness.'" (EX2036, 1 (emphasis added).)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

297. The independent claims of the Challenged Patents do not claim removal of liveness data. In fact, FaceTec has at least one other patent application, US 18/627,421, directed to this unclaimed feature. The application claims "[a] method of validating an item, having a graphic storing code, is associated with a person" wherein the method comprises evaluating three-dimensionality, and if the person is determined to be three-dimensional, then "processing one or more of the at least one first image, the at least one second image, the first data and the second data to generate a face derived code, the face derived code representing the person's face but is not capable of being reverse processed to generate the image of the person from which it was derived." (EX1080, Claim 1.)

298. Therefore, removal of the liveness data is a critical feature, according to FaceTec, of ZoOm to prevent against, for example, honeypot risk. It is a feature that is unclaimed by the independent claims reviewed by Dr. Turk with respect to nexus.

299. In sum, the FaceTec's product, ZoOm, has at least five critical features, according to FaceTec, that are unclaimed by the challenged independent claims: 1) control of the camera, 2) the UI with ovals, 3) encryption of data, 4) liveness checks, and 5) removal of liveness data. Therefore, the product is not coextensive with the challenged independent claims in my opinion.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

B. FaceTec's Evidence Lacks Nexus In Fact To The Claimed Invention

300. I understand that, even if nexus not presumed, FaceTec may still be afforded the chance to prove nexus. However, I also understand that FaceTec must show that the purported objective evidence of nonobviousness is the direct result of the unique characteristics of the claimed invention.

301. As I described above, I understand that FaceTec considers ZoOm to be directed to determining the three-dimensionality of the user and that the independent claims are directed to determining the three-dimensionality of a user or his or her face. (*See* §8.A.) After reviewing the evidence presented by FaceTec, I believe that the objective evidence lacks a nexus to the purported invention of the challenged independent claims and/or fails establish what FaceTec asserts it proves.

1. Alleged praise lacks nexus to challenged claims, and Jumio was not a competitor

302. I understand that evidence of industry praise, including that of a competitor, can be indicia of nonobviousness. However, I also understand that such praise must have a nexus to the claimed invention.

303. I understand that FaceTec points to several certifications as evidence of praise. (POR, 7-8.) I also understand that all of these testing services, iBeta,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

BixeLab, and Praetorian, are services that FaceTec paid to perform the tests and receive the certification. (EX1082, 107:12-17.)

304. I understand that a FaceTec product received an iBeta Level 1 certification. (EX2037, 1; EX2078, 1.) As a first matter, the product that received the certification was the overall ZoOm product (for both Level 1 and Level 2 testing). (EX2037, 1, EX2078, 1.) I understand that Dr. Turk equates the ZoOm product to the portion of the video of Mr. Rose [redacted] (EX1059, 82:10-16, 84:19-85:2.) As I previously discussed, the ZoOm product includes many other anti-spoofing measures besides just a three-dimensionality check. In fact, in the article FaceTec cites announcing the iBeta Level 1 certification, it mentions that ZoOm, in addition to a 3D check, uses "[a]dvanced AI [to] determin[e] if concurrent human liveness traits are present." (EX2036, 3.) The "liveness detection method . . . includes texture, reflectivity, movement, 3D depth and much more." (EX2036, 3.) These statements refer to what I previously discussed as liveness checks. I understand that there was nothing specifically listed in the certification that tested whether the three-dimensionality check, on its own, received this certification.

305. It is important to also understand what these certification services are testing, something Dr. Turk admits he was unfamiliar with, despite touting its importance. (EX1059, 171:23-172:4.) iBeta, according to their website, tests in

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

compliance with the ISO/IEC 30107-3 standard and runs Presentation Attack Detection (PAD) testing services. (*See, generally,* EX1068.) There are different levels of PAD testing, each with different lengths of time, expertise of attackers, and types of artefacts. (EX1068.) PAD Level 1 tests with individuals with no expertise using equipment that "is readily available in a normal home or office environment." (EX1068.)

306. Many other companies besides FaceTec have achieved such a certification, such as Aware, HID Global M-Series, ID R&D, Idemia (for both their passive and active liveness), IDmission, ImageWare, Incode, Innovatrics, iProov, OCR Labs, Redrock, Thales, and Zoloz. (EX1068.) Though FaceTec was the first to receive the certification, I understand that FaceTec has noted that "[r]elying on outdated standards and testing criteria, *such as iBeta/ISO 30107-3*, creates a false sense of security. Attack vectors have evolved so quickly that standards released in 2017 were rendered obsolete within just a few years. ISO 30107-3 and iBeta testing does not include new attack vectors . . . [including] two of the most scalable types of attack vectors with the potential to cause widespread damage." (EX1065, 1 (emphasis added).) Therefore, it is my understanding that even FaceTec gives little weight to iBeta and other ISO 30107-3 testing.

307. I understand that a FaceTec product—one that Dr. Turk was unfamiliar with—also received an iBeta Level 2 PAD certification. (EX2037, 1.)

Level 2 testing uses more sophisticated artefacts, including a "3D printer, resin mask, [and] latex mask." (EX1068, 7.) For the specific Level 2 test performed on FaceTec's product, "6 artefacts or species used in the test effort were two life-like dolls, two realistic latex masks, one custom-made resin mask and one realistic silicone mask with a cost of up to $300 each." (EX2037, 1.) As I previously discussed, a three-dimensionality check on its own does not necessarily defend against attacks performed by masks, as the masked users are still three-dimensional. Therefore, I understand that the Level 2 testing has no nexus to the claimed three-dimensionality checks of the independent claims.

308. I understand that a FaceTec product received a BixeLab Level 2 PAD certification. (EX2038, 1.) FaceTec cites to this certification as additional support for the idea that FaceTec's software received numerous certifications. (POR, 7.) However, the Level 2 PAD, as mentioned above, relates to testing using sophisticated masks, something that a check of three-dimensionality alone does not defend against. Additionally, I understand that many other companies have achieved PAD Level 2 testing according to the BixeLab website, such as OCR Labs, Eperix Inc., BRUK.ID, Digibank, HyperVerge Technologies Private Limited, TNG Digital, Amadeus, FPT IS, Digent ID, Cubertech JSC Vietnam, AwareID, and Oz Forensics Software Trading LLC. (EX1086/EX1070, 1-5.)

309. FaceTec also cites to a second BixeLab certification as evidence. (POR, 7.) However, this certification evaluated the "BRYK.ID Extension v9.6.30 – match-3d-2d-face-portrait API," and measured "Match Rate" rather than protection by different attack vectors. (EX2039, 1.) This 3D face matching technology was offered separately from the 3D liveness check product, as pointed out by Mr. Rose. (EX2059, ¶48; *see also* EX1086/EX1062, 1.) Customers could choose to do only the 3D liveness checks or both the 3D liveness checks and the 3D Face Matching, according to FaceTec. (EX2068, 1.)

Requested Features

Biometric Features Required *

○ 3D Liveness Checks & 3D Face Matching (Note: All Face Matching is Free, Unlimited Usage)

○ 3D Liveness Checks Only (No Face Matching, even though it's Free)

(EX2068, 1.) Therefore, this certification lacks a nexus to the "invention" of the challenged independent claims.

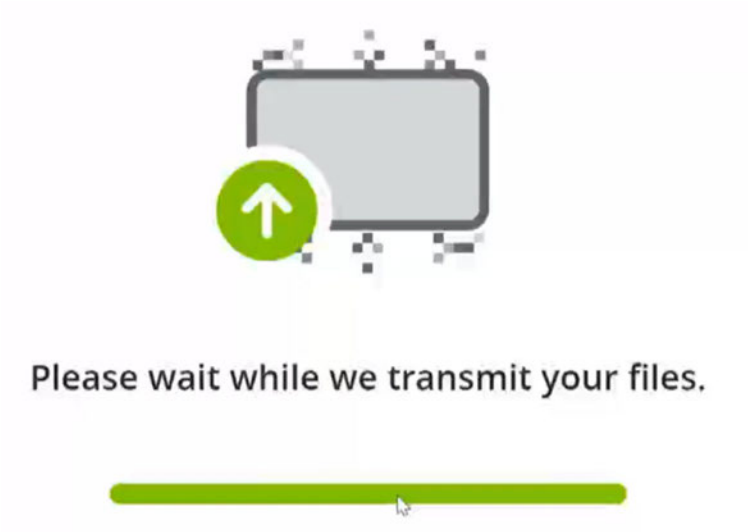
310. I understand that a FaceTec product—again one not familiar to Dr. Turk—received Praetorian's penetration testing certification. (EX2040, 1.) However, Praetorian's certification, the Application Penetration Testing, does not test for liveness. (*See* EX1069.) Instead, according to the Praetorian website, the test "identif[ies] vulnerabilities ranging from classical web application vulnerabilities on the OWASP Top 10 to complex attack paths that chain together multiple weaknesses." (EX1069, 1.) The testers use "[m]anual techniques, such as code review, log analysis, and analyzing memory in a debugger provide depth of

coverage" as well as use "[a]utomated techniques, such as fuzzers and custom enumeration scripts." (EX1069)

311. Such attacks go more to *cybersecurity* than *biometric security*. (EX1082, 127:2-18.) As I previously discussed, such attacks would not be prevented or tested by verifying three-dimensionality. Rather, it would be other features, such as encryption of the data, sending to a secured server, and removal of liveness data that would help with cybersecurity-like issues, all of which are features beyond the scope of the independent claims.

312. FaceTec also points to statements made by Jumio. I understand that, at the time the statements were made, Jumio was *not* a competitor of FaceTec. (*See* EX2057, 11:10.) Rather, as FaceTec pointed out, Jumio was a partner and licensee at the time. (POR, 16; EX2057, 11:10.) Additionally, the product and features that Jumio discusses lack a nexus to the purported invention of the challenged independent claims. For example, EX2057, a talk by Jumio, discusses their partnership with FaceTec and a demo of the product. The product shown and discussed is ZoOm, which, as I previously discussed, includes several other features beyond the three-dimensionality check, such as sending to a secure network shown below.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109



(EX2057, 12:03.) [redacted] Mr. Rose has emphasized that the Face Matching feature is a separate capability from the 3D liveness checks. (EX2059, ¶48.) The talk FaceTec offered as evidence of praise was directed to the combination of *both* the 3D liveness check as well as Face Matching. (*See, generally,* EX2027.) Therefore, the talk shown in EX2057 lacks a nexus to the "invention" of the challenged independent claims.

313. According to FaceTec, article EX2079 highlights how ZoOm is "secure," "accurate," and has an "intuitive interface." (POR, 8.) As mentioned above, ZoOm includes more than just the check of whether a person is three-dimensional, and I understand that the comments in EX2079 are directed to at least some of these other features. (*See* §8.A.1-5.) For example, ZoOm is "secure" at least partially due to the encryption of data and camera control, as I previously

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

discussed. (EX2079, 2; *see* §8.A.1, §8.A.3.) As another example, ZoOm is "accurate" and has "anti-spoofing capabilities" because of its ability to detect other types of attacks in addition to the three-dimensionality check, such as using the liveness checks or preventing deepfake injections. (EX2079, 2; EX1073, 13:20-22; *see* §8.A.1, §8.A.5.)

314. According to FaceTec, a Jumio article, EX2042, makes note of the testing that FaceTec's product underwent and the certification it achieved. (POR, 15.) However, as I previously discussed, FaceTec themselves described these tests as "outdated." (EX1065, 1; *see* §8.B.1.) Additionally, those testing methods included "using live human test subjects wearing 3D masks and deepfakes," attacks that three-dimensionality checks alone are not designed to defend against. (EX2042, 3; *see* §8.B.1.)

315. According to FaceTec, EX2044 and EX2045 show marketing materials featuring "FaceTec's 3D liveness detection technology" including the use of ovals for processing. (POR, 8.) However, as I previously discussed, the ovals have to do with the UI, which is an unclaimed feature with its own, separate patents. (*See* §8.A.2 (note, however, that I make no comment as to the validity of the other patents and patent applications cited or referenced in this declaration other than the four Challenged Patents).) Additionally, the screenshot of the marketing materials shows no praise for the "3D selfie technology" – it only

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

acknowledges it as a product offered. (EX2044, 2.) Rather, the explanation of capabilities is listed separately under the "Face Match" product description, something that FaceTec has admitted is a separate product. (EX2044, 2; EX2059, ¶48.)

2. No proof iProov or Jumio copied the claimed invention

316. I understand that evidence of copying can be indicia of nonobviousness. However, I also understand that mere allegations of copying are insufficient.

a. iProov

317. I understand that FaceTec and Dr. Turk allege that iProov copied the claimed invention. (EX2061, ¶465.) However, I understand FaceTec has not put forth any proof establishing that iProov actually copied the claimed invention.

318. Whether access is established or not, in my opinion, FaceTec put forth very little or no evidence of copying and failed to establish a nexus between what it claims was purportedly copied and the "invention" of the challenged independent claims. Mr. Rose, when discussing copying, only broadly notes that iProov's program verifies that an online user is "[a] real person" and not a "photo, a mask, or other presentation attack," and that it copied "at least two security features" – the UI [redacted] (EX2059, ¶46.) [redacted]

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

[redacted] (*See* §8.A.1-2.) Furthermore, Dr. Turk admitted that liveness detection can include many different things and is not limited to evaluating three-dimensionality. (EX1069, 78:3-8.) Determining that a person is "real" can occur in ways other than three-dimensionality in the independent claims of the challenged patents. (EX1069, 77:24-78:8.)

319. The only other evidence FaceTec offers takes the form of a single paragraph in Dr. Turk's declaration in which he cites to a series of appendices. (*See* POR, 15-16; EX2061, ¶468.) I understand that Dr. Turk admits that he neither prepared the appendices nor reviewed the software of the iProov product, something he admits he would like to do to better to understand the iProov product. (EX1059, 181:5-8, 186:25-187:4, 188:24-189:6, 109:6-11.)

320. In the appendices Dr. Turk cites to, he only relies on two exhibits to demonstrate what the iProov product is doing – EX2071 and EX2080. (*See* EX2061, Appendices A-D.) In my opinion, neither of these exhibits provide any evidence as to the operation of the iProov product.

321. For example, in Appendix C for element 1[d] of the '938 patent, Dr. Turk points to a single screenshot of EX2071 and notes that "[t]he image . . . shows iProov's system allowing user to log in based on the biometric profile, which indicates that iProov authenticates using the biometric data obtained by processing the second image." (EX2061, Appendix C.) However, claim element

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

1[d] recites: "*determine* that the *three-dimensional shape* is not exhibited when the second biometric data does not *have expected differences compared to the first biometric data*, . . . wherein the expected differences result from *fish-eye type distortion* in at least one of the at least one first image and the at least one second image and due to the three-dimensional nature of the human face and the change in distance between the camera and the user." (EX2061, Appendix C.) Nothing in the screenshot or the rest of the video of EX2071 shows any kind of determination of three-dimensionality resulting from expected differences between images. When questioned, Dr. Turk admitted that he "[did]n't know for sure" whether "biometric information was actually compared" and that he would need to "look more closely at iProov's product to determine that." (EX1059, 186:25-187:2.)

322. He also admitted the same, when looking at element 1[d6] of claim 1 of the '471 patent ("comparing the first data to the second data to determine whether expected differences exist between the first data and the second data which indicated three-dimensionality of the user") and EX2071. He testified that he "[did]n't believe the image itself shows that step" of actually "comparing." (EX2061, Appendix A; EX1059, 186:25-186:4.)

323. Additionally, FaceTec did not even allege that iProov specifically made the determination step discussed above. Rather, FaceTec says that, for the iProov product, "[t]he iProov user interface presents an oval and forces the

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

changing the displayed area of the background while leaving the oval size the same. While FaceTec's technology changes the size of the oval and leaves the displayed area of the background the same, the result is the same: the user is forced to change their distance from the camera to complete the session. Accordingly, iProov's software is substantially similar to FaceTec's patented software." (POR, 15-16.) FaceTec makes no allegation that iProov determines if a user is three-dimensional based on expected differences between two images. (POR, 15-16.) Therefore, I understand that FaceTec has not put forth any evidence that iProov copied the claimed invention.

b. Petitioner

324. I understand that FaceTec and Dr. Turk allege that Jumio copied the purported invention of the challenged independent claims. (EX2061, ¶465.) However, I understand FaceTec has not put forth any proof establishing that Jumio actually copied.

325. FaceTec, again, put forth very little or no evidence of copying and failed to establish a nexus between what it claims was purportedly copied. The only evidence FaceTec offers takes the form of four paragraphs in Dr. Turk's declaration in which he cites to a series of appendices. (*See* POR, 16-17; EX2061, ¶470-73.)

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

326. Dr. Turk admits that he neither prepared the appendices nor reviewed the software of the Jumio product. (EX1059, 181:5-8, 189:22-190:2, 109:6-11.) In the appendices Dr. Turk cites to, he relies on only a few exhibits to demonstrate what the Jumio product is doing. (*See* EX2061, Appendices E-H.) In my opinion, none of these exhibits provides any evidence as to the operation of the Jumio product.

327. For example, in Appendix G, element 1[d] of the '938 patent, Dr. Turk points to two screenshots from EX2070 and cites to EX2074 and EX2076. (EX2061, Appendix G.) However, element 1[d] recites: "determine that the three-dimensional shape is not exhibited when the second biometric data does not have expected differences compared to the first biometric data, . . . wherein the expected differences result from fish-eye type distortion in at least one of the at least one first image and the at least one second image and due to the three-dimensional nature of the human face and the change in distance between the camera and the user." (EX2061, Appendix G.) Nothing in the screenshots or the two cited references show any kind of determination of three-dimensionality resulting from expected differences between images.

328. When questioned, Dr. Turk admitted that he was only making an inference and "ha[d] not seen [Jumio's] software and been able to analyze exactly how their system works." (EX1059, 189:17-190:2.) Furthermore, Dr. Turk cites a

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

Jumio article in support of the idea that the element is practiced, noting that Jumio's "system compares the images and automatically provides scores on validity/liveness and similarity." (EX2061, Appendix G.) However, the article discloses nothing about a check for three-dimensionality based on expected differences between images, and Dr. Turk has admitted that there is more than one way to evaluate liveness. (EX2076, 2; EX1059, 78:3-8.)

329. Additionally, FaceTec did not even allege that Jumio specifically made the determination step discussed above. Rather, FaceTec says that Jumio's product "provides a user interface that takes an image of the user at a first distance, instructs the user to 'move closer,' and takes an image of the user at a second distance. It then processes the images." (POR, 16.) FaceTec makes no allegation that Jumio determines if a user is three-dimensional based on expected differences between two images. (POR, 16.) Therefore, I understand that FaceTec has not put forth any evidence that Jumio copied the claimed invention.

3. Claimed invention did not, alone, satisfy the long felt need, and there is no evidence of failure by others

330. I understand that evidence of a long-felt need and failure by others can be indicia of nonobviousness. However, I also understand that such a need must actually be satisfied by the claimed invention, and that there must be evidence that others actually tried and failed if failure is being asserted.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

331. I understand that FaceTec is referring to a need for "software-based 3D liveness detection." (POR, 18.) I understand that none of the evidence that FaceTec cites specifically mention liveness based on "three-dimensionality." Others in the industry did meet the need for a "secure software-based liveness solution" by using other approaches, such as eyeball tracking or monitoring facial movements. (*See* EX2048, EX2049.) I understand too that there is no evidence that these approaches actually failed.

332. I understand that FaceTec approached the need by utilizing perspective distortion to determine whether the user was three-dimensional. (POR, 25.) FaceTec's software then used the determination of three-dimensionality as a proxy for liveness. (POR, 6.) However, as I previously discussed, determining liveness based solely on three-dimensionality is incomplete, as it can be easily spoofed. (*See, e.g.,* §8.A.4.) Thus, FaceTec employs a "multilayered" approach to determining liveness, such that FaceTec's ZoOm is not limited to just determining liveness based on three-dimensionality. (*See* §8.A, EX2059, ¶49.) FaceTec even admits that "true liveness detection requires wholistic observation of the characteristics of the face and the surrounding environment." (EX2036, 3.)

333. Indeed, the features identified as an unclaimed, critical feature, according to FaceTec, are layers for securely determining liveness beyond three-dimensionality. For example, the camera control serves to, among other things,

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

prevent deepfake injections that could spoof the system on liveness. (*See* §8.A.1.) FaceTec itself has called deepfakes and video injections "two of the most scalable types of attack vectors with the potential to cause widespread damage." (EX1065, 1.) As another example, the encryption of the image data (the FaceScan), and its subsequent decryption and processing over a secured server, are needed to prevent a bad actor from taking over the data before it reaches the server and spoofing the server on liveness or potentially stealing the data (i.e., middle-man attacks). (*See* §8.A.3.) As another example, liveness checks are needed to prevent other types of spoofs, such as those tested at PAD Level 2 like three-dimensional masked attackers. (*See* §8.A.4, §8.B.1.) Finally, the deletion of liveness data is needed to prevent a honeypot risk – something that could spoof the server in a subsequent authentication on liveness. (*See* §8.A.5.)

334. FaceTec also points to its certifications again as evidence of their claimed invention supposedly satisfying the need. (POR, 20.) However, as I previously mentioned, the certifications lack a nexus to the "invention" of the challenged independent claims. (*See* §8.B.1.) Additionally, FaceTec itself has called such certifications "outdated" and "create[s] a false sense of security." (EX1065, 1.)

335. FaceTec also alleges that Jumio tried and failed to meet the aforementioned need. (POR, 18.) However, if FaceTec defines the need as

"Secure Software-Based ***3D*** Liveness Detection Technology," FaceTec has put forth no proof that Jumio or anyone else attempted to solve a problem related to evaluating liveness based on three-dimensionality. (POR, 18; *see* §8.B.2.) Additionally, I understand that Petitioner used the term "upgraded," but an "upgrade" does not necessarily mean its previous approach failed or that they abandoned previous approaches. (*See* Ex. 2079, 2.)

4. Industry was not skeptical of the claimed invention

336. I understand that evidence of industry skepticism can be indicia of nonobviousness. However, I also understand that such skepticism must have a nexus to the claimed invention and that a lack of enthusiasm does not equate to skepticism.

337. I understand that FaceTec cites a single article it claims was skeptical of the "invention" of the challenged independent claims. (POR, 21.) Dr. Turk also cites the article, alleging that the article "questions whether such a solution [as the one provided by FaceTec] could succeed." (EX2061, ¶482.) The article does indeed discuss how Apple took a hardware-based approach to determine if the user is human. (EX2050, 1.) However, the article merely contrasts the approach taken by FaceTec as compared with Apple, calling it "a different tactic," rather than using any language of skepticism. (EX2050, 1.) In fact, the article notes that such a technology may even help "FaceTec gain a toehold in the market by making the

feature seem more desirable," and that several banks were already conducting pilot tests with it. (EX2050, 2.)

338. Dr. Turk claims that the "[e]xperts cited in the article voiced doubts." (EX2061, ¶483.) FaceTec notes that one of the people quoted in the article, Marios Savvides, was skeptical if FaceTec's product "could prevent someone with a mask from accessing the handset." (POR, 22.) However, as I previously discussed, the three-dimensionality check alone cannot defend against masked attacks. (*See* §8.A.4.) Further, Savvides explicitly mentions he is *not* skeptical, describing how "FaceTec's method makes sense, and it should help deflect some spoofing attacks like those using photographs in an attempt to get into your phone." (EX2050, 2.)

339. I understand that FaceTec and Dr. Turk make a similar remark about another quoted individual, Rich Mogull, who noted that "being totally software-based means 'it would need additional hardware protection to achieve equivalent levels of security' to Apple's face-scanning technology." (POR, 22; EX2061, ¶483.) FaceTec also notes that Mogull said he would be "shocked if [the FaceTec technology] was as usable or secure as alternative methods we now see being adopted." (POR, 23.) However, FaceTec and Dr. Turk ignore a key part of Mogull's commentary. Mogull added, as mentioned in the article, that "FaceTec can *surely* get its technology to work . . . and it would be more secure than simple two-dimensional facial recognition." (EX2050, 3 (emphasis added).) Noting that

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

something can "surely" work is not indicative of skepticism, and its comparison to alternative methods expresses a lack of enthusiasm, at best.

340. Lastly, FaceTec points to the iBeta certifications as evidence of industry skepticism. (POR, 23.) However, as I previously discussed, such iBeta certifications lacked a nexus to the challenged claims. (*See* §8.B.1.) Additionally, as I previously mentioned, FaceTec was themselves skeptical of the iBeta certification, calling it "outdated" and that it doesn't test new attack vectors such as deepfake injections. (*See* §8.B.1; EX1065, 1.)

5. No nexus between alleged commercial success and claimed invention, and FaceTec statements are contradictory

341. I understand that evidence of commercial success can be indicia of nonobviousness. However, I also understand that such commercial success must have a nexus to the claimed invention.

342. I understand that Mr. Rose points to receiving the certifications from iBeta and others as evidence of commercial success. (EX2059, ¶47.) As a first matter, receiving a certification for the operation of a product is not an indicator of *commercial* success. Further, as mentioned above, those certifications were for the ZoOm product which lacks a nexus to the claimed invention. (*See* §8.A.)

343. Mr. Rose also points to the fact that customers only paid for the 3D liveness check product and not other features, such as Face Matching, and that they

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

used FaceTec "solely for the performance of the 3D liveness checks." (EX2059, ¶¶48-49.) Mr. Rose also cites to alleged increases in use and sales of these 3D liveness check products as well as number of customers. (*Id.*, ¶¶51-53.) FaceTec claims that its success isn't attributable to other "customer-facing features." (POR, 26.) However, as I previously discussed, the 3D liveness check product, ZoOm, itself has many other features besides the three-dimensionality check that is claimed by the challenged independent claims, some of which may not necessarily be "customer-facing." (*See* §8.A.) FaceTec itself calls the 3D liveness check "multilayered" and that the check for three-dimensionality is only the first layer. (POR, 25.) For example, ZoOm also includes the deepfake injection prevention, liveness checks, and encryption of the data, to name a few. (POR, 25.) FaceTec argues that the "claimed invention is driving demand," but Jay Meier noted that the removal of the liveness data in particular was a "critical selling feature in our view." (POR, 25; EX1073, 10:3-4.) Therefore, I understand that there is no nexus between the purported invention of the challenged independent claims and any of the commercial success evidence provided.

344. Mr. Rose also argues that Jumio's license of the software from FaceTec is evidence of commercial success. (EX2059, ¶50.) [redacted]

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

[redacted] Jumio paid FaceTec for each of these services. Therefore, the cost and numbers associated with that license lack nexus.

345. FaceTec also points to evidence of Jumio's license of iProov's software as evidence of commercial success. (POR, 25.) First, as I previously discussed, FaceTec has not put forth any evidence that iProov actually copied or practices the "invention" of the challenged independent claims. (*See* §8.B.2.i.)

[redacted]

346. Lastly, FaceTec claims it was not the market incumbent. (POR, 26.) However, I understand that FaceTec, in that same paper, claims that it "was the first of its kind and effectively created the market for secure software-based 3D liveness detection." (POR, 23.) FaceTec also calls its own product a "disruptive technology and breakthrough product." (POR, 24.) It is my understanding that these statements are contradictory and should therefore be afforded little weight.

Jumio EX1060
Jumio v. FaceTec IPR2025-00109

* * *

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and that these statements were made with knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code.

Dated: Jan 21, 2026

Dr. Christopher Daft

Jumio EX1060
Jumio v. FaceTec IPR2025-00109